STUDIES ON THE FLAVIN-MEDIATED CYSTEINE SALVAGE PATHWAY, STRUCTURAL ENZYMOLOGY OF RIBOFLAVIN LYASE AND MECHANISTIC INVESTIGATION OF THE RADICAL S-ADENOSYL-L-METHIONINE-MEDIATED TRYPTOPHAN LYASE

A Dissertation

by

DHANANJAY BHANDARI

Submitted to the Office of Graduate and Professional Studies of
Texas A&M University
in partial fulfillment of the requirements for the degree of

DOCTOR OF PHILOSOPHY

Chair of Committee,	Tadhg P. Begley
Committee Members,	Frank M. Raushel
	David P. Barondeau
	Paul D. Straight
Head of Department,	Simon W. North

December 2017

Major Subject: Chemistry

ABSTRACT

This dissertation discusses mechanistic and structural studies of enzymes with unique features. The studies are focused on: (i) mechanistic investigation of a novel cysteine salvage pathway that involves a flavoenzyme-mediated Pummerer-type rearrangement; (ii) structural enzymology of the superoxide radical-mediated catabolism of riboflavin; and (iii) mechanistic studies on a radical-mediated fragmentation-recombination reaction catalyzed by Tryptophan Lyase (NosL).

Amino acid cysteine and its metabolites play a vital role in all forms of life. Recently, a pathway was identified in *Bacillus subtilis* that salvages cysteine from S-substituted cysteines. We have reconstituted all five enzymes involved in this transformation. Detailed biochemical studies were performed on a unique Pummerer-type rearrangement catalyzed by a flavin monooxygenase, CmoJ. Though the pathway shows broad substrate specificity, we have identified the physiological substrates of the enzymes. This study provides insights into a novel pathway that extracts sulfur from the organosulfur molecules in the largely unexplored world of rhizosphere microbiome.

In contrast to the biosynthesis of cofactors, their catabolism is scantly described in literature. The first riboflavin catabolic pathway was recently identified in Begley lab in 2013. Mechanistic investigation using biochemical studies revealed that flavin mononucleotide (FMN) dependent Riboflavin Lyase (RcaE) utilizes a superoxide radical to catabolize riboflavin. Crystal structure of RcaE was solved at 1.75 Å, bound with

cofactor FMN and substrate riboflavin. Orientations of these ligands in the active site demonstrate that stacking of the riboflavin on top of FMN prevents the formation of typically observed C_{4a}-flavin(hydro)peroxide. This structural study provides insights into a new class of flavin-dependent enzymes.

Tryptophan Lyase (NosL) is a radical SAM enzyme that catalyzes the formation of 3-methyl-2-indolic acid from L-tryptophan. As suggested by the recent crystal structure of NosL, we have obtained biochemical evidence for H-atom abstraction from the amino group of L-tryptophan. Further experiments demonstrate that the subsequent β-scission involves C_{α}-C(O) cleavage. A mechanistic proposal is discussed that incorporates all latest discoveries, including the newly identified byproducts of the NosL catalyzed reaction. Additional studies with substrate analogs as well as mutants demonstrate a cornucopia of reactions with the 5′-deoxyadenosyl radical, which initiates chemistry in all radical SAM enzymes.

कर्मण्येवाधिकारस्ते मा फलेषु कदाचन ।

Karmanye vadhikaraste Ma Phaleshu Kadachana.

To work you have the right, but not to the fruits thereof.

ACKNOWLEDGEMENTS

I feel entitled to use this section as an opportunity to thank all the people who have played a role in my journey. Their faith, guidance and support have been the pillars of my research endeavors. This is probably the only platform where I can collectively express gratitude to all individuals who have contributed scientifically and otherwise.

First and foremost, I thank my graduate advisor, Prof. Tadhg Begley, for giving me the opportunity as well as the freedom to explore numerous ideas. Tadhg introduced me to the research in biological chemistry and infused it with the excitement of making discoveries. His philosophy and approach towards science, to explore the unknown and willingness to tackle some of the most complex problems in enzymology have left a lasting impression on me. I am indebted to Tadhg for instilling the art of formulating scientific thoughts and communicating them effectively.

I express my deepest gratitude to Prof. Frank Raushel, Prof. David Barondeau and Prof. Paul Straight, for serving on my graduate committee. I am grateful for their insightful feedback during my seminar, presentations and preliminary examinations. I appreciate the efforts of Prof. Pingwei Li and Baoyu Zhao for training me on each and every aspect of protein crystallography, from setting up screens to depositing structures. Their enthusiasm and encouragement motivated me to pursue this challenge tirelessly.

Thanks to Swapnil, Angad and Siddhesh who were instrumental in my decision to enroll for graduate studies at Texas A&M University. I am thankful to Dinuka, Hui and Dinesh for being my mentors in early days of graduate school. Thank you, BJ for all

the late night discussions on projects and Lisa for constant encouragement. Prem and Lina for being great batchmates with whom I have shared numerous responsibilities.

I am grateful for the contributions of all current and previous Begley lab members, especially Kalyan, as I learned various techniques on setting up enzymatic reactions through his lab notes. Thank you Bekir for training me on kinetics and Rung-Yi for teaching peptide/proteomic analysis. Sameh for guidance on synthesis, Dmytro for energetically synthesizing molecules that were key in studying the NosL catalyzed reaction and Yindrila for the preliminary work on RcaE crystallography. Yuanyou, Isita, Sanjoy, Jian and Nitai for their enthusiasm in science. I am thankful to Sumedh for being a great roommate and for the countless discussions we have had almost every day for past three years. Most of our experiments can be traced to these deliberations.

I am indebted to the Department of Chemistry and Prof. Tadhg Begley for giving me several opportunities to present my research to visiting Professors and host lunch sessions with fellow graduate students. These interactions with leading experts were tremendous learning opportunities. I am also grateful to the Department and University for travel awards that gave me opportunities to attend conferences. I thank Dow Chemical and GRC committee for honoring my research efforts with prestigious awards.

I am extremely thankful to several people in the Department of Chemistry at Texas A&M University who impeccably handled all our administrative work. I am grateful to Sandra Horton, Judy Ludwig, Prof. Joanna Pellois and Prof. Simon North for their support. Eva for efficiently handling all group related activities, Ed for ensuring the smooth functioning of ILSB and Viktor for his assistance in fixing instruments. Phillip,

Melvin, Divina and Sherry in the Chemistry & Bio-Bio stockrooms for admirably handling most of our orders and supplies. I also admire the efforts of the staff scientists in several departmental facilities. Yohhanes Rezenom for his help in optimizing GC-MS analysis and Doyong Kim for protein analysis on nano-ESI. Howard Williams for training on NMR instruments and also for his help in acquiring NMR spectra of key low yield molecules on cryo-probe NMR. I am also grateful to Jeremy Wood for his assistance in collecting X-ray data from APS.

I would also like to thank Akshay, Ankush, Tejas, Shamik, Satyajit, Santosh, Himanshu, Piyush, Vinay, Ravi, Pratik, Sagar, Ameya and Tejas for cherishable memories in Cherry 201, College Station. My friends from ICT Abhiram, Sarang, Sharvari, Soumyadeep, Abhay, Tushar, Nikhil, Shweta, Vaishali, Sumit, Rishabh and Onkar for continuing the companionship and backing me up. My school buddy Ashwin, for sharing his experiences while on an expedition to become a “real” doctor.

Throughout this journey, I have been immensely fortunate to have the support of family. My late grandparents, who infused the importance of overcoming obstacles to pursue one’s dreams. My parents, for giving me the exposure, shaping my core values and principles as I grew up. My siblings, Ganesh and Pradnya, for their consistent encouragement and filling me up on all family events that I missed over the last few years. Finally, I am indebted to Shweta Kide and her family for their infallible love, patience, blessings and trust. This dissertation would not have been possible without the foundations laid by everyone mentioned above. Thanks to all my friends, colleagues, faculty and staff for making my time at Texas A&M University a great experience.

CONTRIBUTORS AND FUNDING SOURCES

This work was supervised by a dissertation committee consisting of Professor Tadhg P. Begley (advisor), Professor Frank M. Raushel and Professor David P. Barondeau from the Department of Chemistry and Professor Paul D. Straight from the Department of Biochemistry and Biophysics.

Preliminary work on the enzymes discussed in sections 2.3.1 and 2.3.2 was performed by Dr. Kalyanraman Krishnamoorthy. Screening of organisms that catabolize riboflavin reported in section 3.2.1 was performed by Dr. Benjamin Philmus while Dr. Hui Xu identified the gene cluster involved in this catabolism (section 3.2.2). Dr. Yindrila Chakrabarty characterized enzymes in the riboflavin catabolic pathway described in Section 3.2.2 and 3.2.3. Protein structures illustrated in section 2.3.6 and 3.3 were obtained in collaboration with Prof. Pingwei Li and assistance of his postdoc Dr. Baoyu Zhao. Mutants of NosL described in section 4.4.6 were cloned by Dr. Hui Xu. Synthesis of compound **115** was carried out by Dr. Dmytro Fedoseyenko. Cryoprobe NMR of compounds **172** and **175** were recorded by Dr. Howard Williams. All other work conducted for the dissertation was performed by the student independently.

Graduate study was supported by funding from following:

1) Graduate Teaching Assistant in Organic Chemistry labs within Chemistry Department at Texas A&M University.
2) Funding from Robert A. Welch Foundation grant A-0034.
3) Funding from National Institutes of Health (NIH) grant DK44083.

NOMENCLATURE

5′tA	5′-deoxy-5′-thioadenosine
Ado$^{\bullet}$	5′-deoxyadenosyl radical
Ado-H	5′-deoxyadeosine
ATP	Adenosine triphosphate
BDE	Bond dissociation energy
BLAST	Basic Local Alignment Search Tool
BSTFA	N,O-Bis(trimethylsilyl)trifluoroacetamide
CIP	Calf intestinal phosphatase
DMSO	Dimethyl sulfoxide
DNA	Deoxyribonucleic acid
DTNB	5,5′-dithio-bis-(2-nitrobenzoic acid) Ellman's reagent
DTT	Dithiothreitol
EIC	Extracted Ion Chromatogram
EPR	Electron Paramagnetic Resonance
FMN	Flavin mononucleotide
Fre	Flavin reductase form *E.coli*
GC-MS	Gas Chromatography-Mass Spectrometry
HPLC	High Performance Liquid Chromatography
kda	kilodalton
KPi	Potassium phosphate

LC-MS	Liquid Chromatography-Mass Spectrometry
MsrA	Methionine sulfoxide reductase A
MTA	Methylthioadenosine
NAD^+	Nicotinamide adenine dinucleotide (oxidized form)
NADH	Nicotinamide adenine dinucleotide (reduced form)
NCBI	National Center for Biotechnology Information
NMR	Nuclear Magnetic Resonance
NosL	Tryptophan Lyase
OD_{600}	Optical density at 600 nm
PDB	Protein Data Bank
PFBHA	o-(2,3,4,5,6-pentafluorobenzyl)-hydroxylamine
PLP	Pyridoxal-5′-phosphate
RBF	Riboflavin
RcaE	Riboflavin Lyase
RNA	Ribonucleic acid
RT	Room temperature
SAM	S-Adenosyl-L-methionine
TCEP	Tris(2-carboxyethyl)phosphine
TFA	Trifluoroacetic acid
THF	Tetrahydrofuran
Tris	Tris(hydroxymethyl)aminomethane

TABLE OF CONTENTS

LIST OF FIGURES

Page

LIST OF TABLES

Page

1. INTRODUCTION

1.1 Introduction to cofactors[1]

Cofactors assist enzymes in catalyzing reactions. They can either be organic or inorganic in nature. They are either bound to the enzyme covalently or held by non-covalent interactions. They are found ubiquitously in all forms of life and play a critical role in several key transformations.[2] A subset of cofactors which we cannot biosynthesize and are supplemented through our diet are called vitamins. Lack of these vitamins can lead to debilitating diseases and even result in death. There has been tremendous progress in the past century to understand the role of cofactors at the molecular level.

1.2 Flavins

Flavins are yellow chromophore found in various food sources like meat, eggs, dairy products, whole grains and dark green-vegetables. Alexander Wynter Blyth originally discovered flavins in milk as a yellow-green pigment in 1872.[3] Further studies in 1934 by Paul Karrer and Richard Kuhn elucidated its structure and the yellow chromophore was named riboflavin (**1**, RBF). As shown in Figure 1.1, the physiologically active forms of flavins that act as cofactors include flavin mononucleotide (**2**, FMN) and flavin adenine dinucleotide (**3**, FAD).[4]

Flavins are versatile cofactors that consist of a three membered 7,8-dimethyl-isoalloxazine moiety, which is central to its activity.[5] They can exist in multiple redox

states, undergo reaction with O_2, form covalent intermediates, transport electrons, etc. Deficiency of riboflavin can manifest in conditions like night blindness, migraines, anemia, fatigue, scaly skin, growth abnormalities in infants, etc.[6]

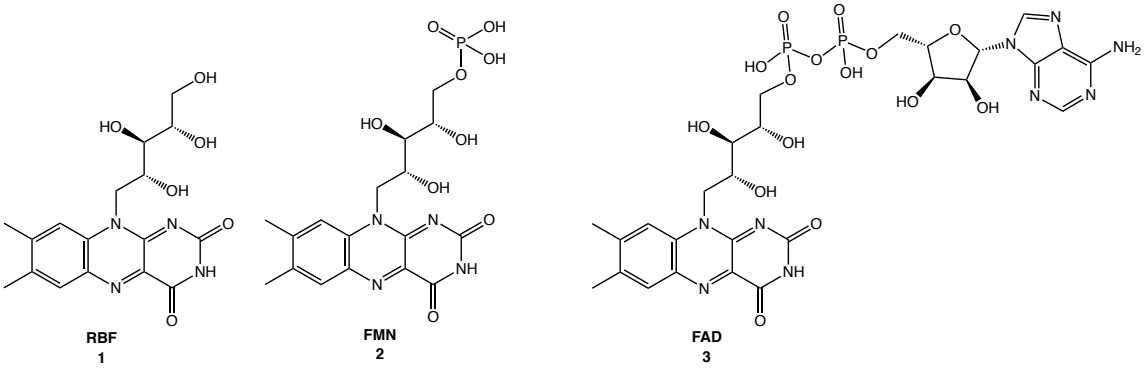


Figure 1.1: Structure of RBF (**1**), its biologically active forms FMN (**2**) and FAD (**3**).

1.3 S-Adenosyl-L-methionine (4, SAM)

S-adenosyl-L-methionine (**4**) is an essential sulfonium compound that is involved in many biochemical processes.[7] It has been studied extensively since its chemical structure was first described in 1951.[8] In many cases, SAM acts as a biological methyl donor, which can affect diverse processes like expression of genes and hormones, fetal development, brain function, etc.[9] In addition to its methyl donor activity, SAM also acts as a substrate in biosynthesis of key compounds such as cyclopropanes, modified tRNA basepairs, N-acylhomoserine lactone, etc.[7]

In 2001, enzymes that utilize [4Fe-4S] and SAM to perform radical mediated transformations were classified into a new enzyme superfamily.[10] Due to exponential

growth in availability of genome sequences in the last decade, this enzyme family currently includes more than 165,000 members.[11] In this superfamily, the first step involves the transfer of an electron from [4Fe-4S] to the sulfonium group of SAM, which leads to the formation 5′-deoxyadenosyl radical (**6,** Figure 1.2A). This highly reactive oxidant typically abstracts a H-atom from the substrates, which further undergo complex transformations to generate products that would be inaccessible without radical chemistry.[12] Figure 1.2B shows the remarkable range of reactions observed with organic radicals, most of which have also been observed with radical SAM enzymes. These include H-atom abstractions, addition to double bonds and β-bond scission reactions.[13]

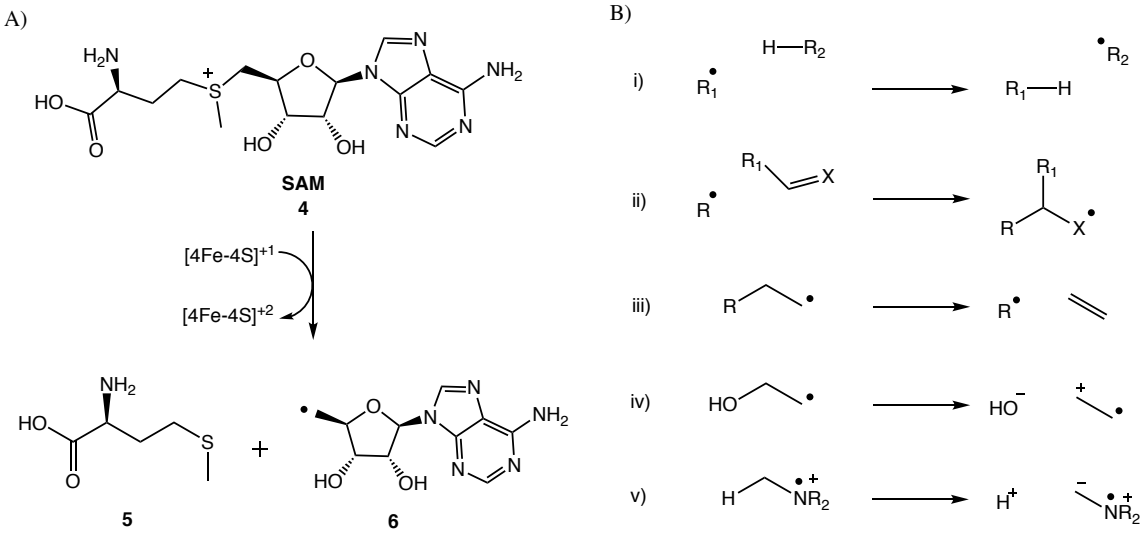


Figure 1.2: Radical mediated chemistry. A) In radical SAM enzymes, reductive homolytic cleavage of **4** generates **6**. B) Examples of major transformations observed in organic radical chemistry. i) Hydrogen atom abstraction. ii) Addition to double bonds. iii-v) β-bond scission reactions.

1.4 Research opportunity

Recently, a pathway was found to be upregulated in *Bacillus subtilis* that salvages cysteine from S-substituted cysteines (**7**) in absence of a direct sulfur source in growth media.[14] The pathway consists of transporter and transmembrane proteins that will help in uptake of organosulfur molecules present in the environment. All previously identified pathways of sulfur assimilation from organosulfur metabolites involve oxidation of sulfur to sulfite, while the unique feature of this pathway is that it liberates L-cysteine (**8**, Figure 1.3). The sulfur from L-cysteine (**8**) can be mobilized to meet the intra cellular needs of sulfur. Interestingly, this pathway generates a reduced form of sulfur even though it consists two flavin dependent mono-oxygenases. Identifying the roles of all enzymes in this pathway will delineate the intermediates and give insights on the strategy evolved by nature to perform a C-S lyase reaction using two flavin mono-oxygenase enzymes.

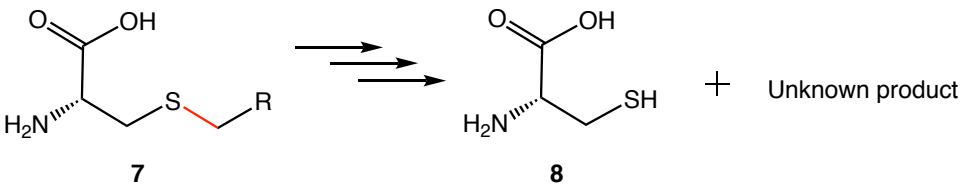


Figure 1.3: Conversion of S-substituted cysteine (**7**) to L-cysteine (**8**) in *B.subtilis*.

Pathways involved in catabolism of riboflavin were reported in literature in the mid 20th century.[15-17] However, most of the strains were lost and enzymes involved in

the transformation were never identified.[18] The first gene cluster involved in catabolism of riboflavin (**1**) was recently identified in the Begley lab.[19] Further biochemical studies on enzymes in the pathway demonstrated that a flavin dependent enzyme (named Riboflavin Lyase, RcaE) catalyzed superoxide radical (**9**) mediated conversion of riboflavin (**1**) to lumichrome (**11**) and ribose (**12,** Figure 1.4). RcaE is the first reported example in flavin enzymology where the superoxide radical (**9**) catalyzed H-atom abstraction from substrate has been thoroughly demonstrated (*manuscript under review*). Structural studies on RcaE will illustrate the unique features this enzyme has evolved to catalyze a new class of flavin-mediated chemistry.

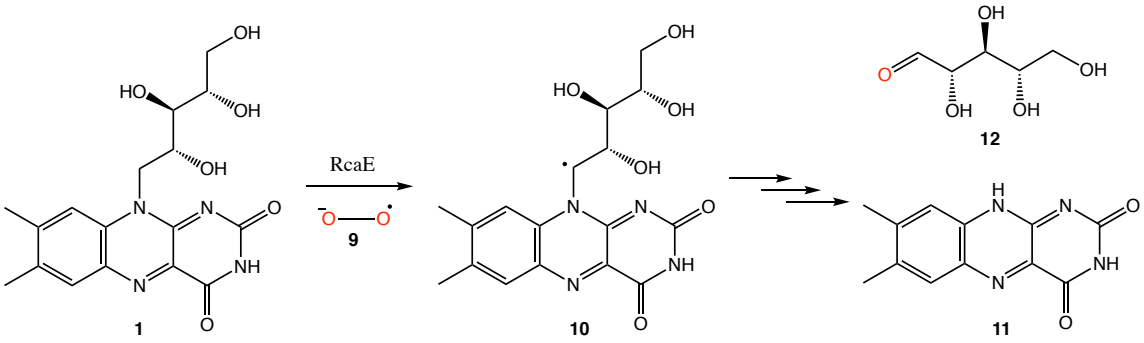


Figure 1.4: RcaE catalyzed superoxide radical (**9**) mediated conversion of riboflavin (**1**) to lumichrome (**11**) and ribose (**12**).

Tryptophan Lyase (NosL) is encoded in the biosynthetic pathway of antibiotic nosiheptide.[20] NosL is a radical SAM enzyme that catalyzes the conversion of L-tryptophan (**13**) to an unusual compound 3-methyl-2-indolic acid (**15**, Figure 1.5).[21] Based on sequence analysis, NosL is classified into the aromatic amino acid lyase

family, which includes enzymes like ThiH, HydG and CofH.[13] Recently published structure of NosL bound to substrate and cofactor suggests that the 5′-deoxadenosyl radical (**6**) abstracts H-atom from the amino group of L-tryptophan (**13**) instead of indolic N-H as previously proposed in the literature.[22] Further experiments would unravel the mechanistic details of the unprecedented fragmentation-recombination reaction catalyzed by NosL.

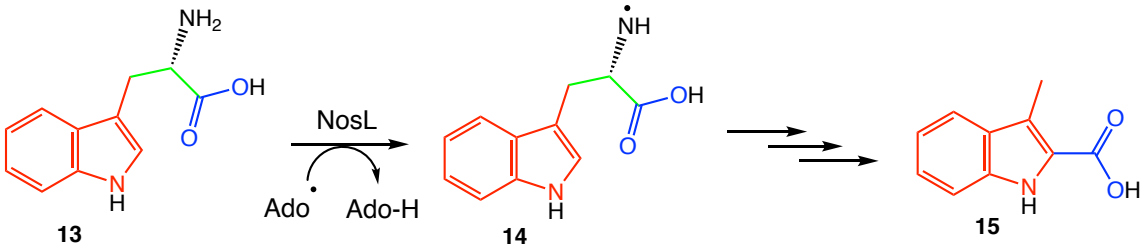


Figure 1.5: NosL catalyzed conversion of L-tryptophan (**13**) to 3-methyl-2-indolic acid (**15**) might progress through H-atom abstraction from amino group of **13**.

2. FLAVIN MEDIATED CYSTEINE SALVAGE PATHWAY

2.1 Role of sulfur in living organisms[23]

Sulfur is the sixth most abundant element in living organisms followed by hydrogen, carbon, oxygen, nitrogen and phosphorus. Sulfur is essential for all living organisms and is required for the formation of many proteins, enzymes, cofactors and other metabolites. It plays a critical role in many catalytic and electrochemical processes in cells. Sulfur containing metabolites like glutathione or bacillithiol function as intracellular reducing agents and prevent damages that can be caused by reactive oxygen species. The high-energy sulfonium metabolite SAM (**4**) serves as a universal methyl donor,[24] while Fe-S clusters play a crucial role in redox chemistry.[25] Sulfur is also a component of life saving antibiotics like penicillin and cephalosporin.[26]

Bacteria obtain sulfur from multiple sources in environment.[27] These include inorganic sources like sulphates, thiosulphates or sulfites as well as organic sources like cysteine, methionine, sulfonates and sulfonate esters. Many organisms have evolved pathways to obtain sulfite (**19**) from organosulfur molecules available in nature like taurine (**16**), alkanesulfonates (**20**) and dibenzothiophene (**22**, Figure 2.1). These pathways are typically upregulated to acquire sulfur from the environment under limiting sulfur conditions. The *tau* operon encodes taurine dioxygenases (TauD) that catalyzes oxygenation of taurine,[28] the *ssu* operon encodes a flavin monooxygenase SsuD that acts on alkanesulfonates[29] and the *dsz* operon encodes a flavin dependent DszA and a hydrolase DszB that catalyze C-S cleavage during dibenzothiophene catabolism.[30]

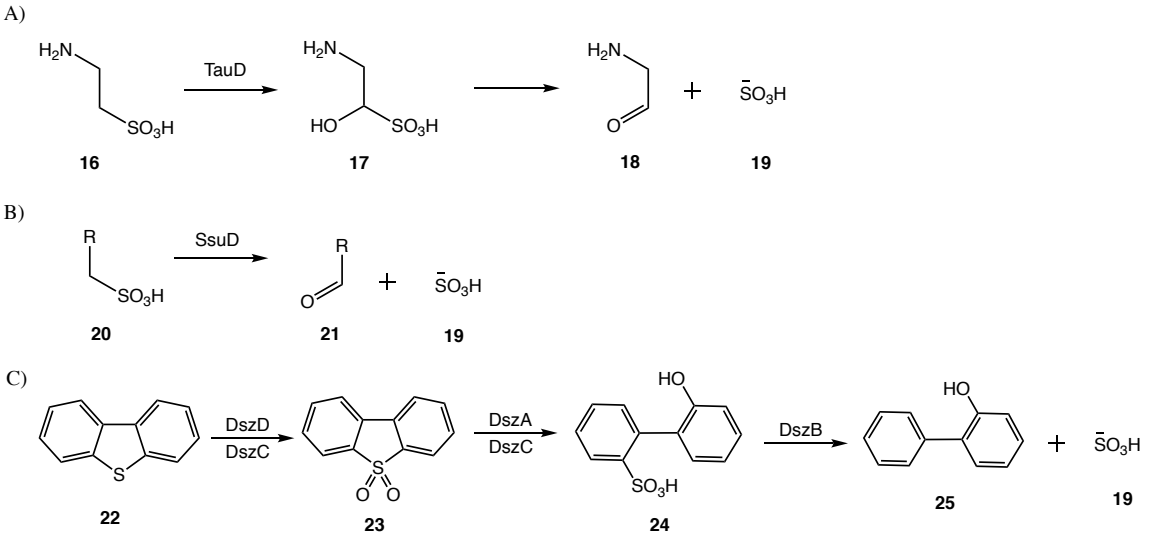


Figure 2.1: Previously reported pathways in which sulfur is retrieved from the organosulfur compounds. A) Taurine (**16**) degradation pathway. B) Alkanesulfonate (**20**) degradation pathway. C) Dibenzothiophene (**22**) desulfurization through 4S pathway.

2.2 Identification of a new pathway involved in S-substituted cysteine (7) catabolism[14]

The gram-positive bacterium *Bacillus subtilis* uses diverse sources of sulphur available in its environment for survival and growth. Recently, growth studies were performed on *B.subtilis* in absence of an inorganic sulfur source but in presence of S-alkylated cysteines shown in Figure 2.2A. Slow growth of bacteria was observed on these organosulfur compounds. This indicates that the bacteria are able to retrieve sulfur from these compounds and there is a novel pathway involved in this transformation. Transcriptome profiling was performed using DNA microarrays to identify the mRNA transcripts that were upregulated in presence of these organosulfur compounds. One of the operons which was upregulated is operon *ytmI_tcyJKLMNytmO_ytnJ_rbfK_ytnLM*

(now renamed *snaAtcyJKLMNcmoO_cmoIJ_rbfK_sndAytnM*) as shown in Figure 2.2B. The putative function of genes in this operon is provided in Table 1.[14]

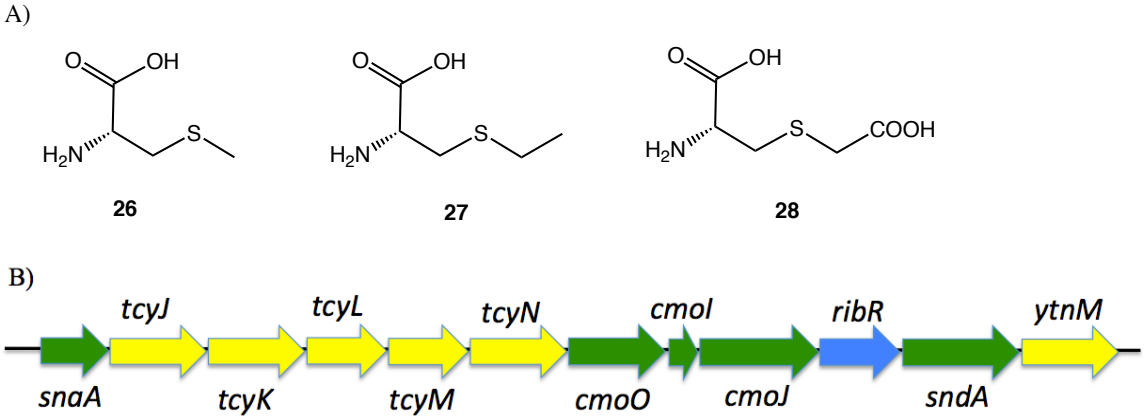


Figure 2.2: Identification of the cysteine salvage pathway. A) S-substituted cysteines used as sulfur source for growth studies in *B.subtilis*. B) Operon that is upregulated when organosulfur compounds are used as a sole source of sulfur in media.

Table 2.1: Functional annotation of genes involved in cysteine salvage pathway.

Sr. no.	Gene	Renamed	Proposed Function
1	*ytmI*	*snaA*	Acetyl transferase
2	*ytmJ*	*tcyJ*	Periplasmic binding protein
3	*ytmK*	*tcyK*	Periplasmic binding protein
4	*ytmL*	*tcyL*	Amino acid ABC-transporter protein
5	*ytmM*	*tcyM*	Amino acid ABC-transporter protein
6	*ytmN*	*tcyN*	Amino acid ABC-transporter protein
7	*ytmO*	*cmoO*	Flavin dependent monooxygenase
8	*ytnI*	*cmoI*	Glutaredoxin
9	*ytnJ*	*cmoJ*	Flavin dependent monooxygenase
10	*rbfK*	*ribR*	Riboflavin kinase
11	*ytnL*	*sndA*	Deacetylase / Amidohydrolase
12	*ytnM*	*ytnM*	Transmembrane protein

The relevance of the genes from this operon in utilizing organosulfur compounds was confirmed by gene deletion studies. Further *in vivo* studies demonstrated that the genes involved in cysteine biosynthesis are not expressed when *B.subtilis* is grown in presence of the organosulfur compounds. This suggested that L-cysteine is salvaged by this pathway instead of obtaining sulfur in the form of sulfite. A model outlining roles of all enzymes in the pathway was proposed as shown in Figure 2.3.[14]

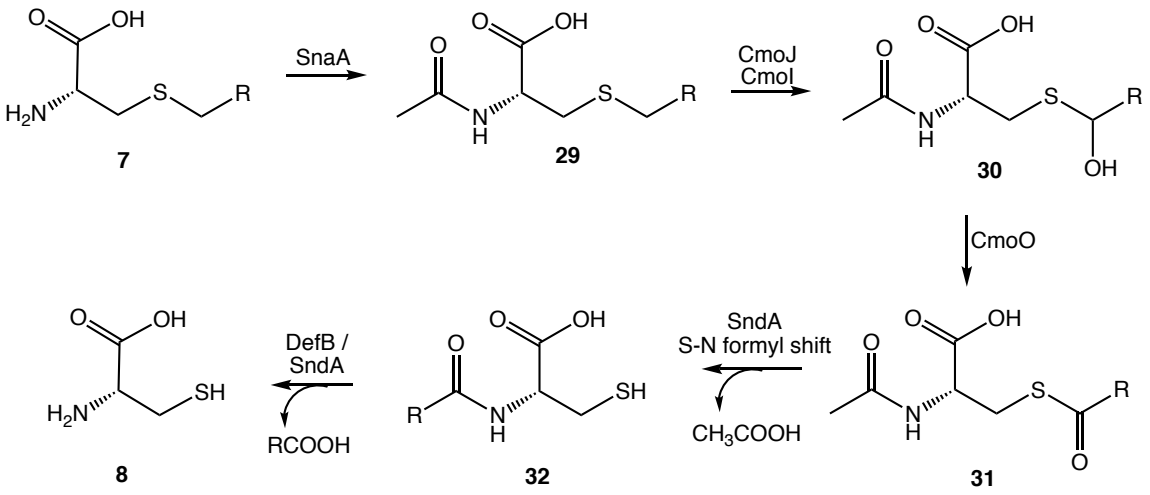


Figure 2.3: Proposed role of each enzyme in the cysteine salvage pathway.

According to the proposal, the pathway is upregulated in absence of direct sulfur sources. The amino acid binding proteins and the periplasmic binding proteins transport organosulfur molecules available in the environment into the bacteria. The S-substituted cysteines (**7**) are then acetylated by SnaA (acetyl transferase) to N-acetyl-S-substituted cysteines (**29**). Action of CmoJ (flavin monooxygenase) and CmoI (glutaredoxin) leads to the formation of hydroxylated compound **30**. This compound is further oxidized to the

corresponding ketone **31** by CmoO (flavin monooxygenase). SndA (amidohydrolase) deacetylates the ketone **31**, which results in the formation of a free amino group. This intermediate undergoes S-N formyl shift to generate compound **32**, which can be deformylated by DefB (deformylase present in *B.subtilis*) or deacetylated using SndA (amidohydrolase) to generate L-cysteine (**8**).

2.3 Results and discussion

2.3.1 Characterization of SnaA (acetyl transferase)

To establish the roles of individual enzymes, we initially focused our attention on SnaA (acetyl transferase) and SndA (amidohydrolase). The activity of purified SnaA was investigated with S-methyl-L-cysteine (**26**), S-carboxymethyl-L-cysteine (**28**) and L-methionine (**5**) in presence of acetyl-CoA (**33**). As shown in Figure 2.4, the assay mixture included Ellman's reagent (**35**), which derivatizes the thiol group of coenzyme A (**34**) and liberates 5-thio-2-nitrobenzoate (**37**) with λ_{max} =412 nm.[31]

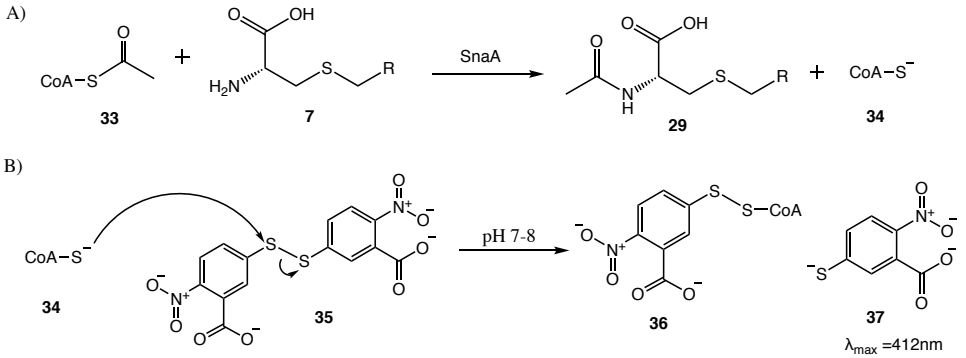


Figure 2.4: Investigating the acetyl transferase activity. A) Transfer of acetyl group to the substrate in presence of SnaA will generate coenzyme A (**34**). B) Derivatization of coenzyme A (**34**) with Ellmans's reagent (**35**) generates chromophore (**37**).

The enzymatic activity was monitored on UV-Vis spectrophotometer at 412 nm. As shown in Figure 2.5, amongst the substrates mentioned above, highest activity was observed with S-methyl-L-cysteine (**26**). The activity of SnaA with L-methionine (**5**) as well as other canonical amino acids was low, which suggests that the enzyme prefers S-substituted cysteines (**7**) as substrates over other amino acids present *in vivo*.

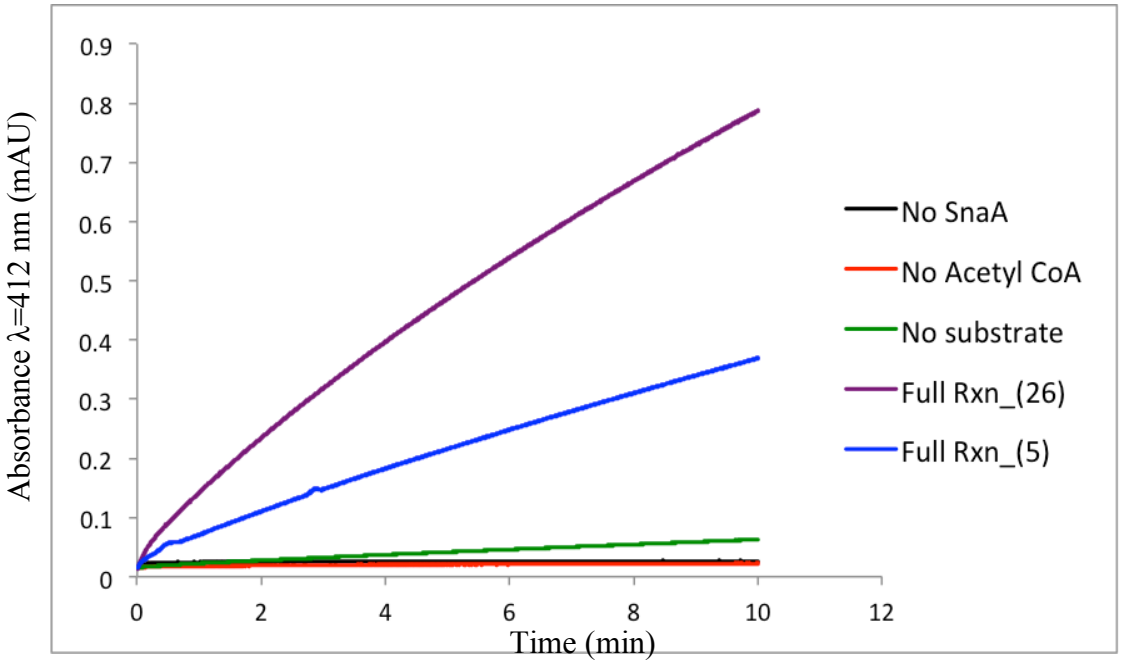


Figure 2.5: Monitoring formation of 5-thio-2-nitrobenzoate (**37**) at 412 nm during SnaA catalyzed reaction.

2.3.2 Characterization of SndA (amidohydrolase)

His-tagged SndA was found to be insoluble during purification since it formed inclusion bodies. The soluble form of the enzyme was obtained by co-expressing it with

tig-chaperone in *E.coli* BL21(DE3) strain.[32] The deacetylase activity was not observed when the enzyme was pre-treated with 8-hydroxy-5-quinolinesulfonic acid to chelate divalent metal ions present during enzymatic reaction.[33] However, the activity was restored by addition of Zn^{+2} or Co^{+2} with substrates like N-acetyl-S-methyl-L-cysteine as well as with N-acetyl-L-cysteine. This suggests that deacetylation might occur during one of the intermediate steps or at the final step in this pathway. The activity was confirmed using LC-MS by derivatizing the free thiol with bromobimane (**40**).[34]

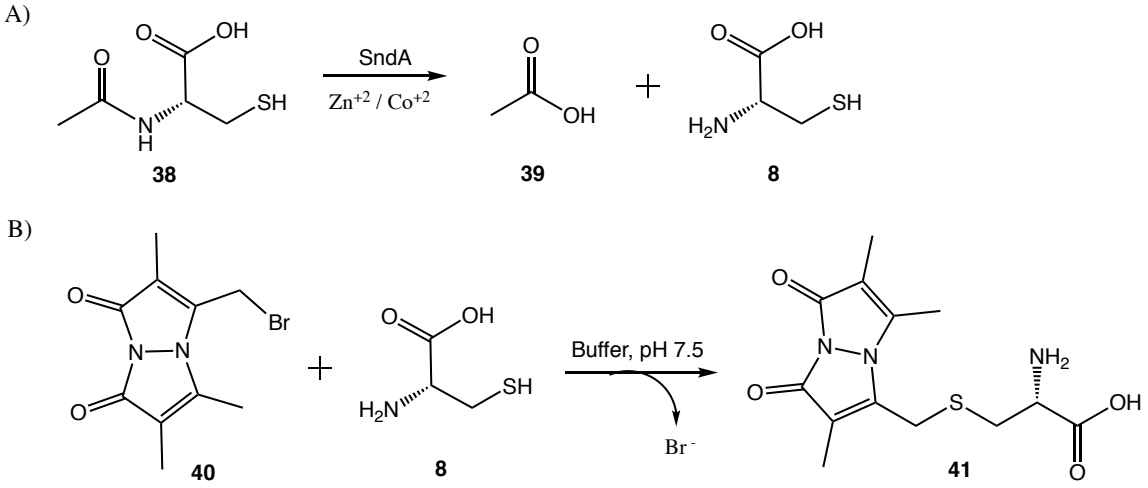


Figure 2.6: Analysis of SndA catalyzed deacetylation. A) Scheme showing reaction catalyzed by SndA. B) Derivatization of L-cysteine (**8**) with bromobimane (**40**). C) LC-MS analysis showing EIC $[M-H]^-$ = 310 Da demonstrating formation of **8** in presence of SndA and Co^{+2}, which is further derivatized with **40**. D) MS of reaction product **41** in negative mode.

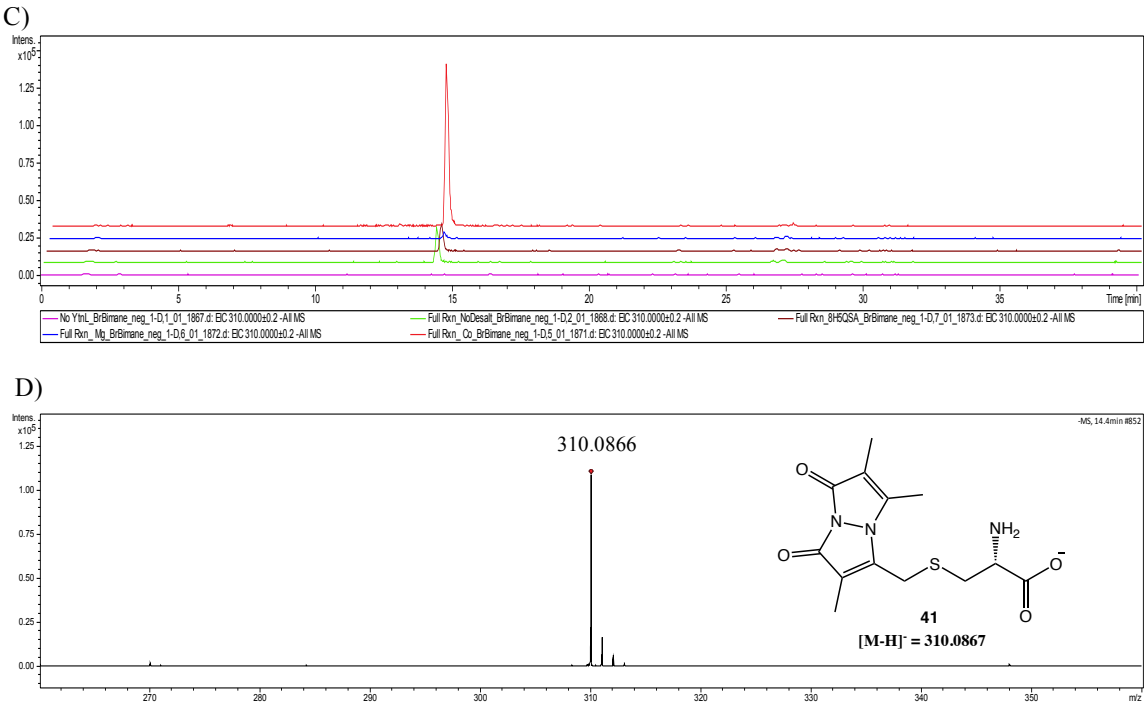


Figure 2.6 Continued

2.3.3 Sulfoxidation catalyzed by CmoO (flavin monooxygenase)

Next, the activity of CmoO, CmoI and CmoJ was investigated with N-acetyl-S-methyl-L-cysteine (**42**) as substrate using ^{1}H-NMR. The enzymes were tested individually as well as in different combinations. In presence of CmoO, a new signal was obtained at 2.7 ppm after ~18 hr. The enzyme utilizes FMN as a cofactor and NADH was added in enzymatic reaction to generate $FMNH_2$. However, bioinformatic analysis and UV-Vis analysis at 340 nm investigating NADH consumption suggested that CmoO lacks flavin reductase activity. To have a continuous supply of $FMNH_2$, flavin reductase (Fre) from *E.coli* was included during the CmoO catalyzed reaction.[35] As shown in

Figure 2.7, within few a minutes disappearance of a signal at 2.1 ppm concomitant with formation of new signals at 2.7 ppm and 3.3 ppm were observed.

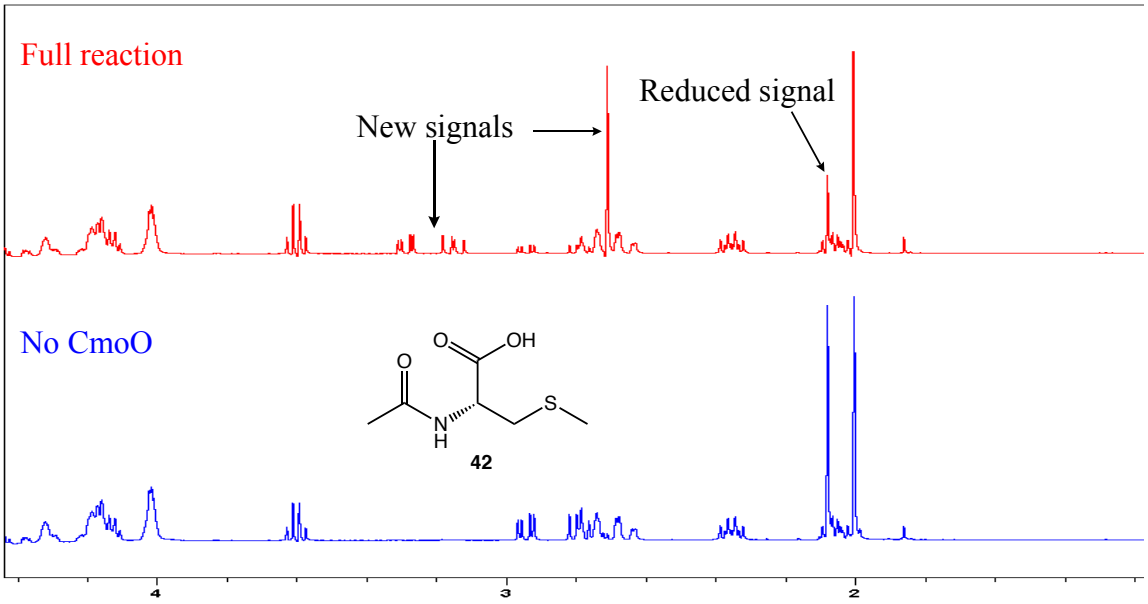


Figure 2.7: ^{1}H-NMR analysis of the CmoO catalyzed reaction with N-acetyl-S-methyl-L-cysteine (**42**) in presence of Fre, FMN and NADH. In 'Full reaction' sample formation of new signals was observed at 2.7 ppm and 3.3 ppm in concert with disappearance of signal at 2.1 ppm.

The downfield shift of protons adjacent to sulfur in N-acetyl-S-methyl-L-cysteine (**42**) suggests that the CmoO has an oxidized sulfur atom. The identity of this product was confirmed to be N-acetyl-S-methyl-L-cysteine sulfoxide (**43**) by spiking enzymatic reaction with the synthesized standard. The synthetic standard of **43** consists of the two diastereomers of the molecule. As shown in Figure 2.8, ^{1}H-NMR analysis at 2.7 ppm

showed that the methyl proton of **43** generated enzymatically corresponds to the methyl protons from one of the diastereomers of **43**.

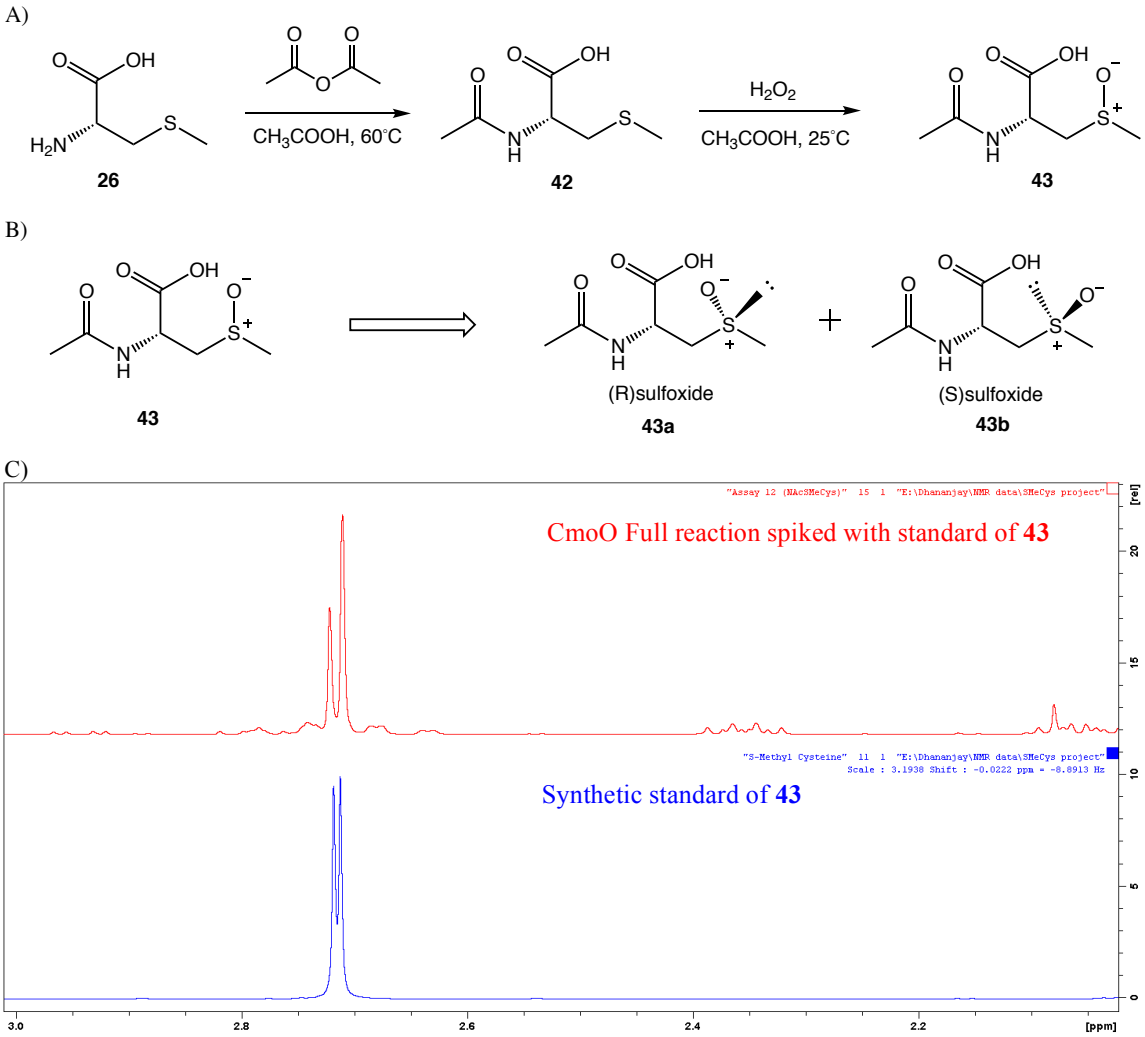


Figure 2.8: Identifying the product generated by CmoO catalyzed reaction. A) Scheme for synthesizing authentic standard of **43**. B) Diastereomers of **43**. C) ^{1}H-NMR analysis of the synthetic standard shows a doublet at 2.7 ppm (blue trace). When the CmoO 'Full Reaction' sample is spiked with **43**, it matches with the upfield proton signal at 2.7 ppm.

2.3.4 Determining the stereochemistry of sulfoxide generated by CmoO

To determine the stereochemistry of **43** generated during the CmoO catalyzed reaction, we inspected enzymes reported in literature that consume stereospecific sulfoxide substrates. Methionine sulfoxide reductases (Msr) are ubiquitously found in many organisms and are involved in repairing oxidative damage to L-methionine (**5**) residue in proteins. MsrA stereospecifically reduces the S-sulfoxide while MsrB reduces the R-sulfoxide stereoisomer.[36] These enzymes show broad substrate specificity and have been previously shown to reduce N-acetyl-S-methyl-L-cysteine sulfoxide (**43**). As shown in Figure 2.9, activity of MsrA affects the downfield signal at 2.7 ppm corresponding to the methyl protons of **43**. This demonstrates that the downfield signal belongs to the S-sulfoxide and upfield signal belongs to the R-sulfoxide stereoisomer.

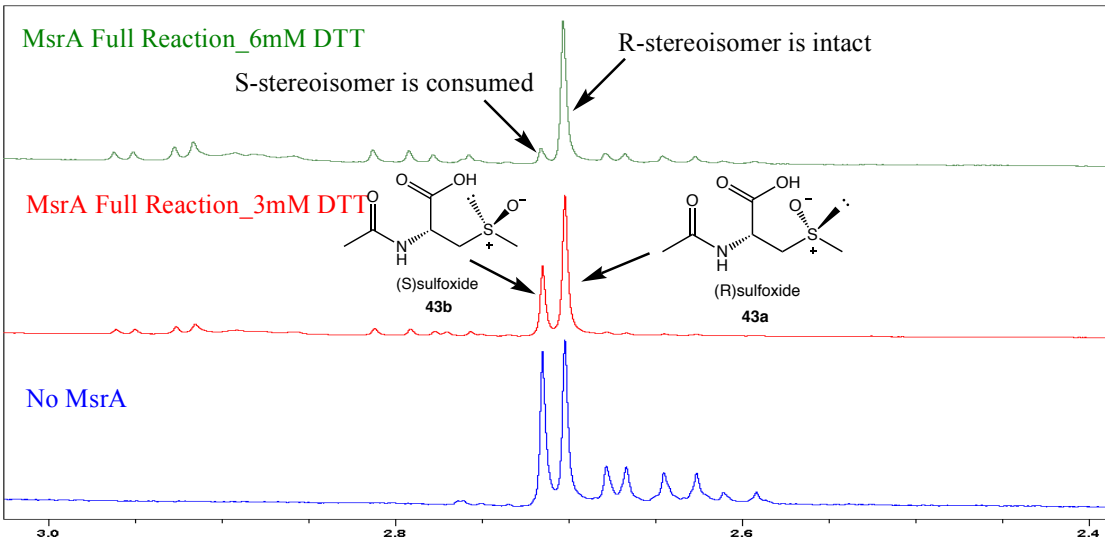


Figure 2.9: Identifying stereoisomers of N-acetyl-S-methyl-L-cysteine sulfoxide (**43**). Analysis of ^{1}H-NMR shows a doublet at 2.7 ppm. When MsrA is added, the downfield signal is consumed which demonstrates that it corresponds to S-stereoisomer of **43**.

In Figure 2.8C, it can be seen that the stereospecific sulfoxide generated from the CmoO catalyzed corresponds to the upfield signal at 2.7 ppm of **43**. By utilizing MsrA, as shown in Figure 2.9, the downfield signal at 2.7 ppm was assigned to the S-stereoisomer whereas the upfield signal corresponds to the R-stereoisomer of **43**. This demonstrates that in the CmoO catalyzed reaction generates the R-stereoisomer of **43**. Additionally, no activity of CmoO was observed with S-methyl-L-cysteine (**26**) while SndA was also unable to deacetylate N-acetyl-S-methyl-L-cysteine-R-sulfoxide (**43a**). Based on the reconstitution and preliminary studies of the above-mentioned three enzymes in this pathway, a revised model was proposed as shown in Figure 2.10.

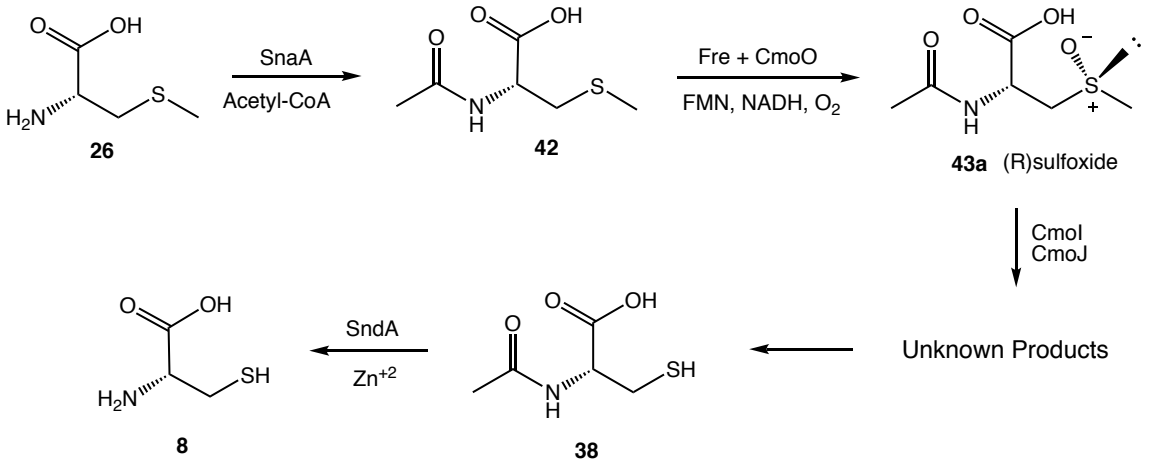


Figure 2.10: Revised model outlining the role of enzymes involved in the cysteine salvage pathway.

2.3.5 Pummerer type rearrangement catalyzed by CmoJ (flavin monooxygenase)

A key question that emerged after reconstitution of CmoO was how a flavin monooxygenase (CmoJ) and glutaredoxin (CmoI) are involved in catalyzing formation of a thiol (**38**) from a sulfoxide (**43a**) in this pathway. The product **43a** of CmoO was tested as a substrate for CmoJ (flavin monooxygenase) and CmoI (glutaredoxin). Despite several attempts of varying *in vitro* enzymatic reaction conditions and performing coupled assays, ^{1}H-NMR analysis did not show consumption of signals corresponding to **43** or appearance of any new signal. With an aim of developing a sensitive assay in which substrate consumption can be monitored using HPLC, I surveyed literature for additional S-substituted cysteines (**7**) that include a chromophore. Molecules like S-benzyl-L-cysteine (**44**) and its sulfoxide, which are observed in roots of some plants, have been previously reported.[37] N-acetyl-S-benzyl-L-cysteine sulfoxide (**46**) was synthesized and tested as a substrate for CmoJ. As shown in Figure 2.11, HPLC analysis showed the consumption of one stereoisomer within few minutes.

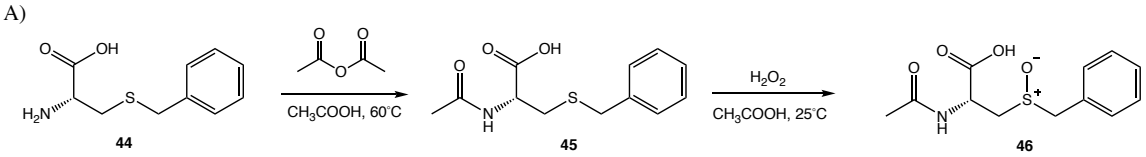


Figure 2.11: Investigating N-acetyl-S-benzyl-L-cysteine sulfoxide (**46**) as a substrate for the CmoJ catalyzed reaction. A) Scheme depicting synthesis of **46**. B) HPLC analysis of the CmoJ catalyzed reaction with **46** (diastereomers) at 220 nm in presence of Fre, FMN and NADH. Consumption of one stereoisomer at 12.5 min was observed in the 'Full reaction' sample.

B)

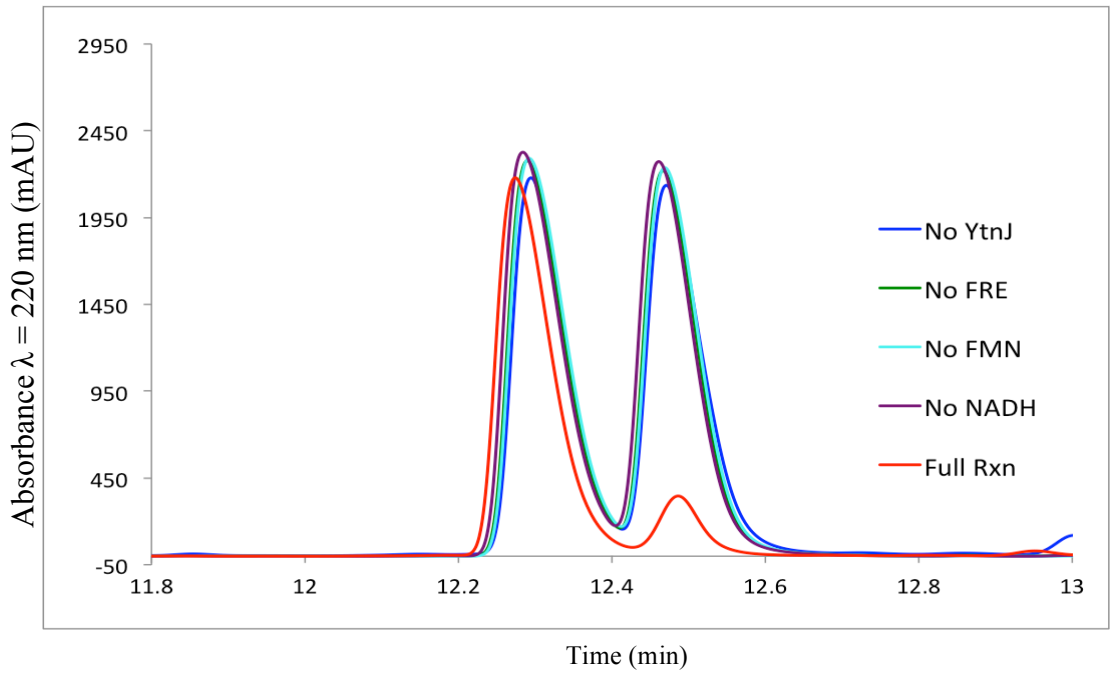


Figure 2.11 continued.

The activity of the CmoJ catalyzed reaction with **46** was quite robust. Further ^{1}H-NMR analysis of enzymatic reaction in the downfield region demonstrated formation of benzaldehyde as one of the products as shown in Figure 2.12. The other product was identified as N-acetyl-L-cysteine (**38**) only in the presence of reducing agents like TCEP.

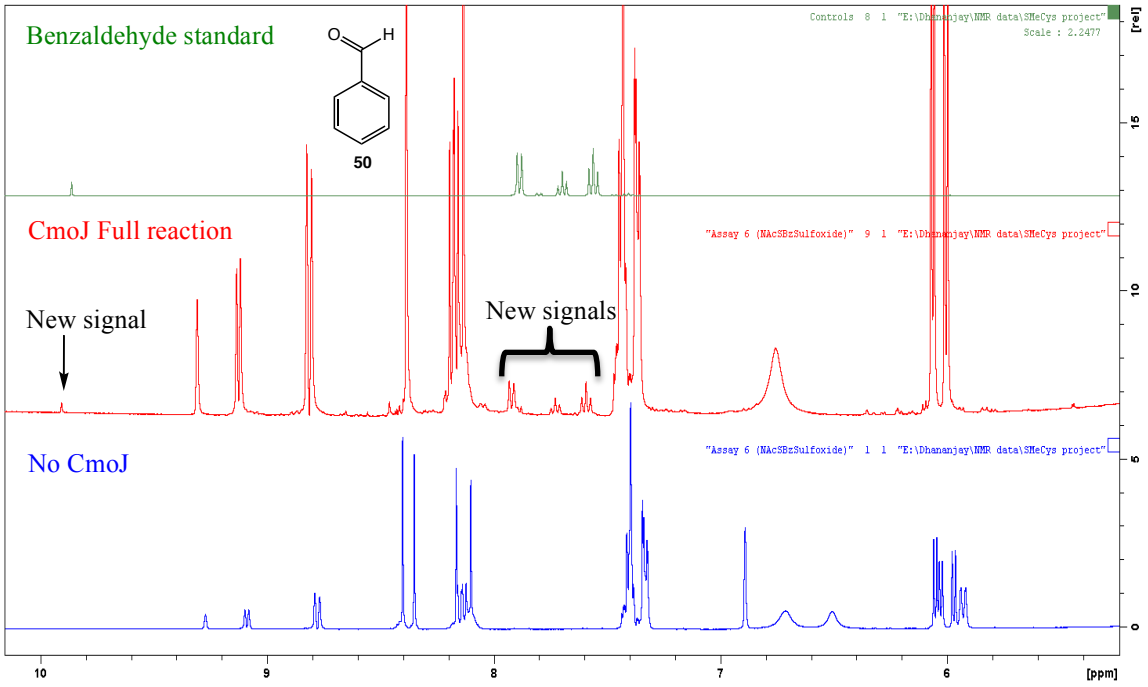


Figure 2.12: ^{1}H-NMR analysis of the CmoJ catalyzed reaction. The downfield region shows formation of new signals in 'Full reaction' sample at 9.8 ppm and between 7.5-8.0 ppm (red trace). NMR of the authentic standard of benzaldehyde confirms that it is one of the products of CmoJ catalyzed reaction with **46**.

A mechanistic proposal accounting for the formation of benzaldehyde (**50**) and an oxidized cysteine containing compound is shown in Figure 2.13A. N-acetyl-S-benzyl-L-cysteine sulfoxide (**46**) is deprotonated at the α-carbon by CmoJ. The resulting carbanion **47** attacks C_{4a}-flavinhydroperoxide, which is a common intermediate in flavin monooxygenases. This leads to the formation of intermediate **48**, which collapses to generate benzaldehyde (**50**) and N-acetyl-L-cysteine sulfenic acid (**49**). This reaction can be classified as Pummerer-type rearrangement. Reduction of sulfonium sulfur

concomitant with oxidation or substitution at the α-carbon is referred to as the Pummerer rearrangement an example of which is shown in Figure 2.13B.[38]

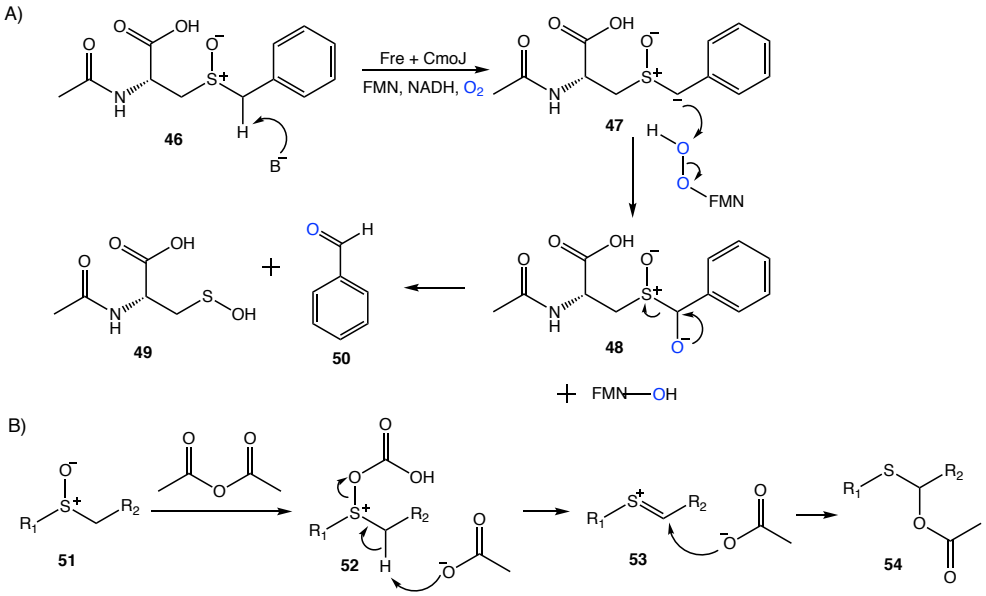


Figure 2.13: Mechanistic proposal of the CmoJ catalyzed reaction with N-acetyl-S-benzyl-L-cysteine sulfoxide (**46**). A) Proposed products with **46** as substrate are benzaldehyde (**50**) and N-acetyl-L-cysteine sulfenic acid (**49**). B) An example of Pummerer rearrangement.

Sulfenic acids are highly reactive species, which can function either as a weak nucleophile or soft electrophile.[39] In absence of any reducing agents, they undergo self-condensation to form thiosulfinates, while they are reduced to thiol in presence of a reducing agent. To investigate the formation of a sulfenic acid **49**, the CmoJ catalyzed reaction was performed in presence of dimedone (**56**).[40] As shown in Figure 2.14, the adduct **57** was detected using LC-MS. To further confirm formation of the sulfenic acid,

the enzymatic reaction was performed in presence of 4-mercaptopyridine (**58**), which derivatizes sulfenic acid **49** and its dimer thiosulfinate **55**.[41] The adduct **59** was established using LC-MS and was further characterized by co-injection with synthetic standard using HPLC and followed by NMR. When the enzymatic reaction was performed in presence of reducing agent like TCEP, formation of N-acetyl-L-cysteine (**38**) was observed in ^{1}H-NMR. This was further confirmed by the derivatization of N-acetyl-L-cysteine (**38**) with NBD-Cl (**60**) and analyzing with HPLC and LC-MS the formation of adduct **61**.

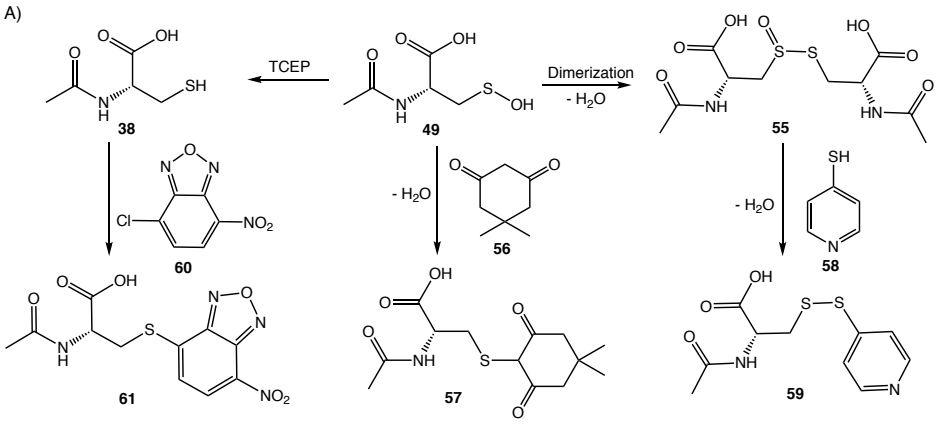


Figure 2.14: Evidence for the formation of sulfenic acid during the CmoJ catalyzed reaction. A) Scheme showing various derivatization reagents used to derivatize sulfenic acid **49**. B) LC-MS analysis showing EIC $[M-H]^-$ = 300 Da demonstrating formation of **57** in presence of dimedone (**56**). The inset shows MS of the reaction product **57** in negative mode. C) LC-MS analysis showing EIC $[M-H]^-$ = 271 Da demonstrating formation of **59** in presence of 4-mercaptopyridine (**58**). The inset shows MS of the reaction product **59** in negative mode. D) HPLC analysis of the reaction mixture treated with **58** at 254 nm showing coinjection with the authentic standard of **59**. E) ^{1}H-NMR and ^{13}C-NMR of the adduct **59**. F) HPLC analysis of the CmoJ reaction treated with NBD-Cl (**60**) at 425 nm confirming formation of **61** only in presence of reducing agents like TCEP. G) MS of the adduct **61.**

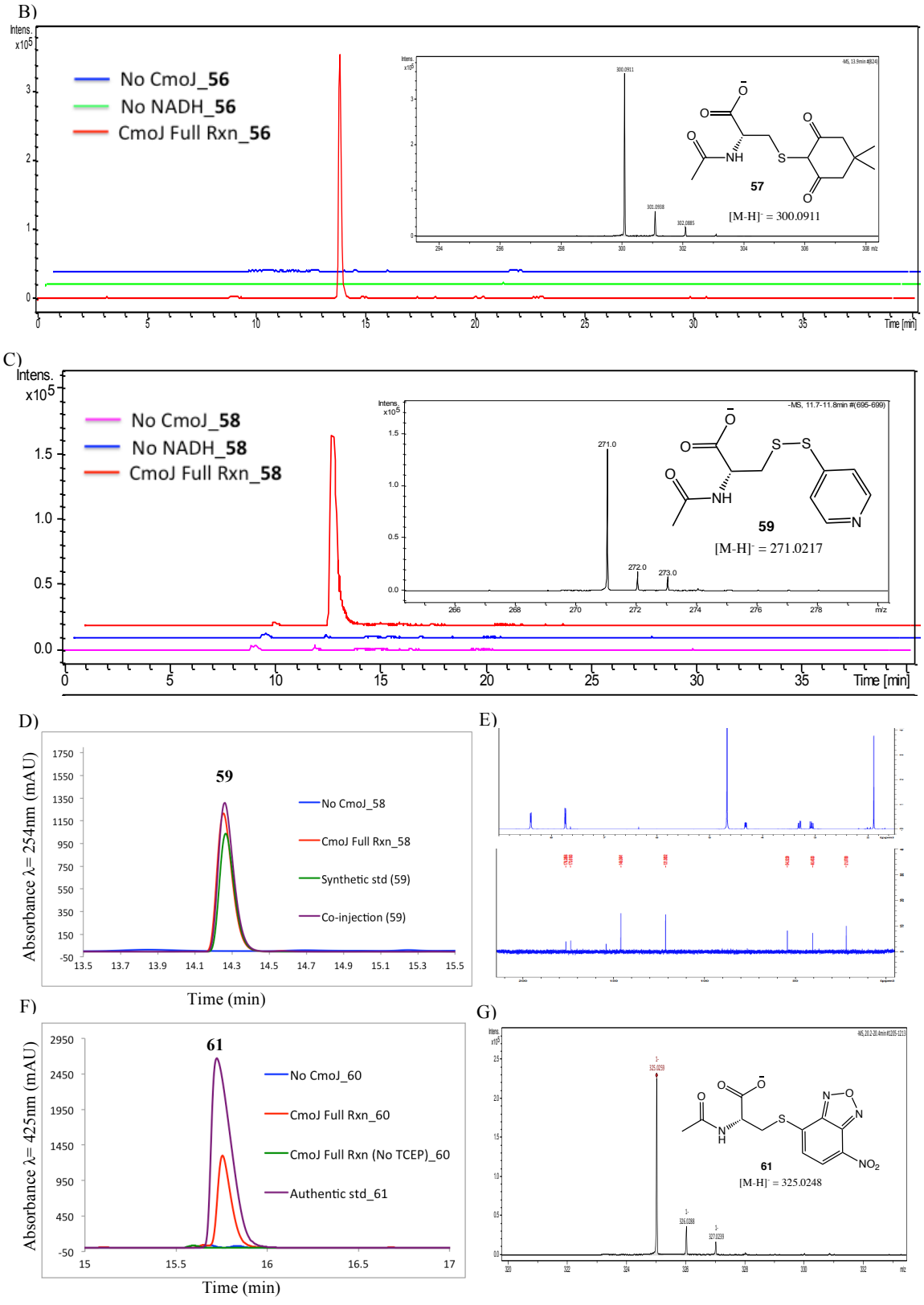


Figure 2.14 continued.

To further investigate mechanistic proposal outlined in Figure 2.13A, the incorporation of O_2 in products can be studied. If enzymatic reaction is performed in presence of $^{18}O_2$, an increase of 2 Da should be observed in benzaldehyde. To decrease the rate of oxygen exchange of benzaldehyde in aqueous environment, an electron-donating group can be added at *ortho* or *para* positions of the benzene ring.[42] Thus, N-acetyl-S-(p-hydroxybenzyl)-L-cysteine sulfoxide (**62**) was synthesized and was found to be a good substrate of CmoJ. The enzymatic reaction was performed in presence of $^{16}O_2$ and $^{18}O_2$, and was quenched after 10 minutes using ethyl acetate to prevent further exchange of oxygen. The organic layer was treated with BSTFA to derivatize the hydroxyl group of **63** and the sample was analyzed using GC-MS.[43] Increase in mass by 2 Da was observed in the trimethylsilyl derivative of p-hydroxybenzaldehyde **64** when the reaction was performed in $^{18}O_2$ confirming it is the source of oxygen in the product.

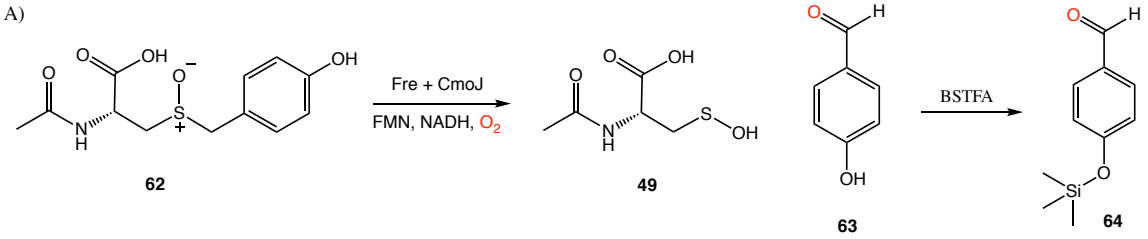


Figure 2.15: Incorporation of O_2 into product during the CmoJ catalyzed reaction. A) Reaction of CmoJ with **62** generates p-hydroxybenzaldehyde (**63**) which is treated with BSTFA to form **64**. B) GC-MS of the product when the reaction is performed in presence of $^{16}O_2$ shows mass = 194 Da. B) GC-MS of the product when the reaction is performed in presence of $^{18}O_2$ shows increase in the mass of product by 2 units to 196 Da.

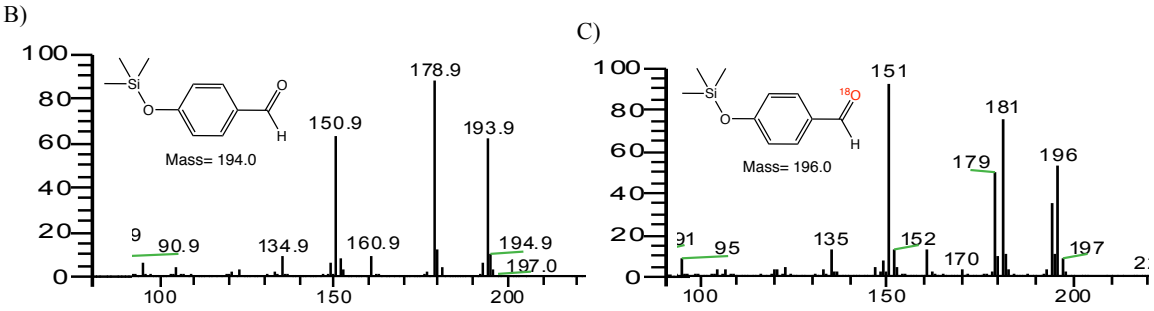


Figure 2.15 continued.

To investigate the substrate tolerance of CmoJ and gain further insights into its substrate specificity, we synthesized variants of S-substituted cysteine (**7**). These analogs were fed to CmoJ in presence of Fre, FMN, NADH and 4-mercaptopyridine (**58**). The formation of adduct **59** was analyzed on HPLC at 254 nm and quantified by generating a calibration curve from the synthesized standard of **59.** Michaelis-Menten kinetic plots were laid out for these analogs and the k_{cat}, K_m values were obtained (Figure 2.16).

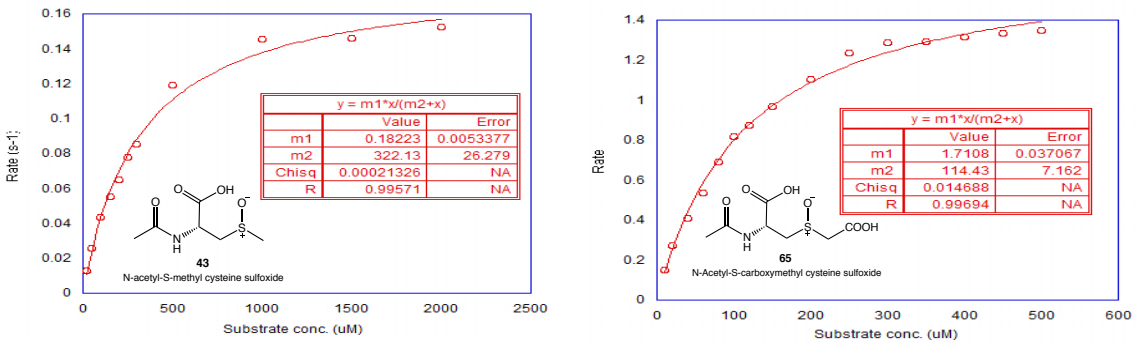


Figure 2.16: Michaelis-Menten kinetic plots of the CmoJ catalyzed reaction with the respective substrate analogs.

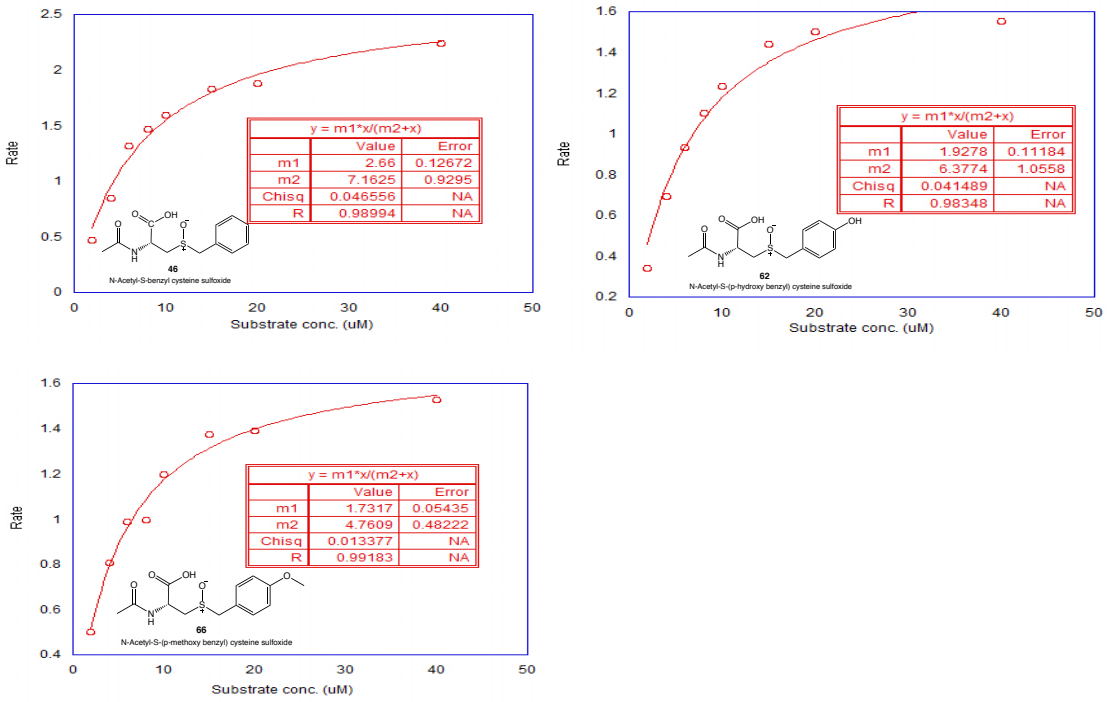


Figure 2.16 continued.

Table 2.2: Summary of k_{cat} and K_m values of various analogs tested with CmoJ.

Sr. #	Substrate analog	k_{cat} (s^{-1})	K_m (μM)	k_{cat}/K_m (M^{-1} s^{-1})
1.	N-Acetyl-S-methylcysteine sulfoxide (**43**)	0.18 ± 0.01	322 ± 26	5.58 x 10^2
2.	N-Acetyl-S-carboxymethylcysteine sulfoxide (**65**)	1.71 ± 0.03	114 ± 7	1.49 x 10^4
3.	N-Acetyl-S-benzylcysteine sulfoxide (**46**)	2.66 ± 0.12	7.1 ± 1	3.72 x 10^5
4.	N-Acetyl-S-(p-hydroxybenzyl)cysteine sulfoxide (**62**)	1.92 ± 0.11	6.4 ± 1	3.63 x 10^5
5.	N-Acetyl-S-(p-methoxybenzyl)cysteine sulfoxide (**66**)	1.73 ± 0.05	4.76 ± 0.5	3.0 x 10^5

As shown in Table 2.2, N-acetyl-S-methyl-L-cysteine sulfoxide (**43**) is a poor substrate for CmoJ followed by N-acetyl-S-carboxymethyl-L-cysteine sulfoxide (**65**), while N-acetyl-S-benzyl-L-cysteine sulfoxide (**46**) was the best substrate amongst all compounds tested. This is coincident with the trend of pKa value of the proton attached to the α-carbon of sulfoxide (deprotonation is the first step in the mechanistic proposal outlined in Figure 2.13A). This also suggests that N-acetyl-S-benzyl-L-cysteine sulfoxide (**46**) might be the native substrate of CmoJ.

2.3.6 Crystal structure of the flavin monooxygenase CmoJ

To gain further insights into the substrate specificity of CmoJ, we turned our attention to the active site architecture of CmoJ and decided to obtain crystal structure of the enzyme. Researchers from the structural genomics consortium have previously deposited the structures of the apo enzyme (PDB id: 1TVL, 2.1 Å) and the FMN bound (PDB id: 1YW1, 2.81 Å) enzyme. However, both the structures have more than 35 amino acid residues (Arg 304 – Glu 341) missing near the FMN binding site and do not offer any information on the residues involved in binding substrate. We obtained the crystals of CmoJ in its apo form (1.69 Å) and also bound with FMN (1.90 Å). As shown in Figure 2.17, the core of the enzyme has a classic TIM barrel fold consisting the $(\alpha\beta)_8$ strands with additional α-helices forming the lid domain (missing in previous structures).

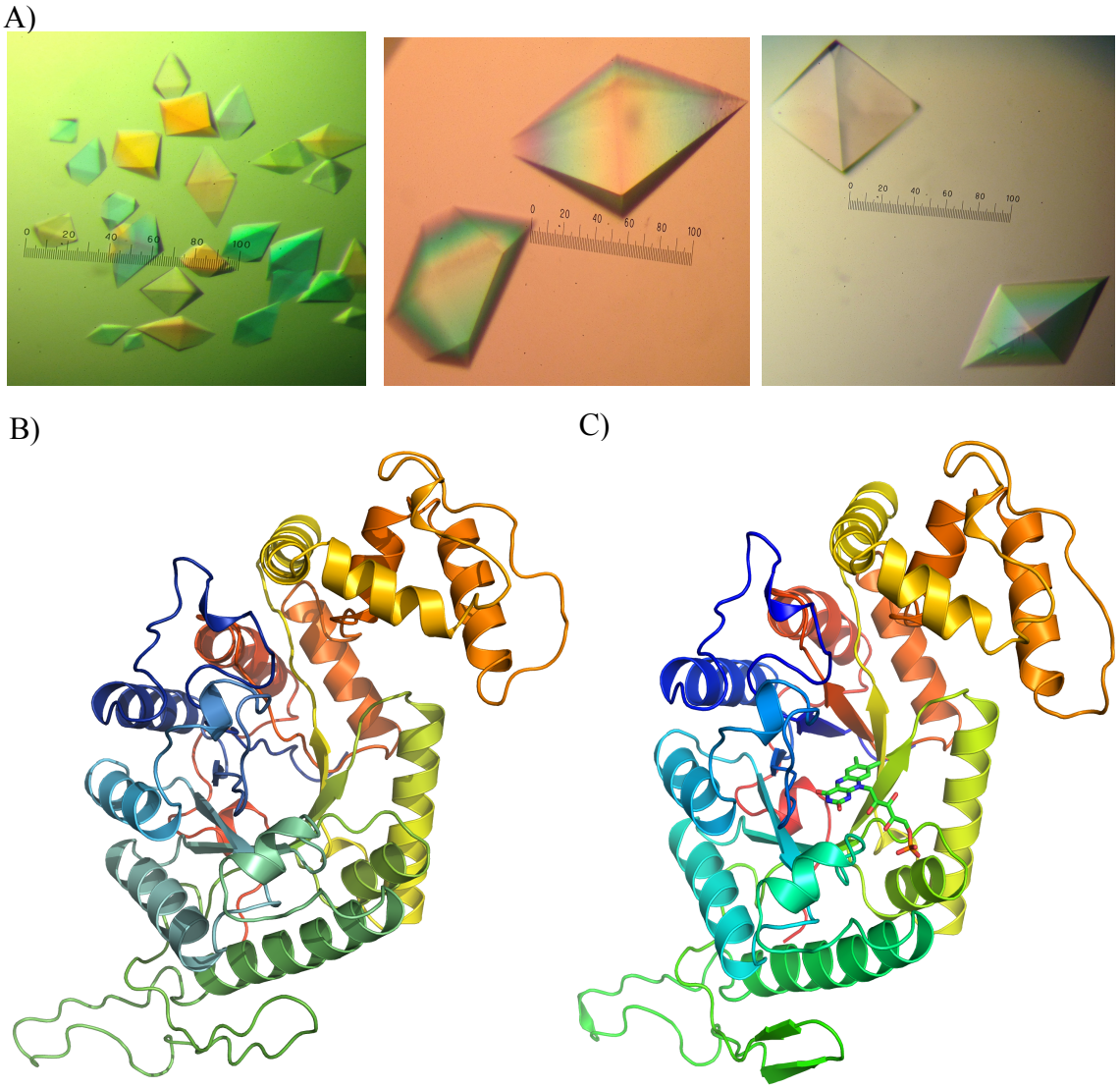


Figure 2.17: Crystal structure of CmoJ. A) Snapshot of the crystals of CmoJ viewed under microscope. B) Crystal structure in the Apo form. C) Crystal structure of the FMN bound CmoJ. A huge empty pocket above FMN can be observed.

Table 2.3: Data collection and refinement statistics for CmoJ.

	CmoJ (Apo form)	CmoJ + FMN + LFN
Data collection		
Space group	P 4_3 2_1 2	P 4_3 2_1 2
Cell dimensions		
a, *b*, *c* (Å)	86.75, 86.75, 195.42	87.07, 87.07, 194.94
α, β, γ (°)	90, 90, 90	90, 90, 90
Resolution (Å)	61.34-1.69 (1.74-1.69) *	38.94-1.90 (1.97-1.90)
R_{merge}	0.14 (1.884)	0.140 (1.029)
$I / \sigma I$	12.1 (1.4)	16.14 (1)
Completeness (%)	100 (100)	100 (99.3)
Redundancy	21.2 (23.4)	6.8 (5.4)
Refinement		
Resolution (Å)	1.69	1.90
No. reflections	84371	59542
R_{work} / R_{free}	0.174/0.190	0.171/0.201
No. atoms		
Protein	3420	3407
Ligand/ion	NA	50
Water	296	216
B-factors		
Protein	37.38	54.23
Ligand/ion	NA	75.47
Water	41.42	49.62
R.m.s. deviations		
Bond lengths (Å)	0.006	0.010
Bond angles (°)	0.750	0.952

*Values in parentheses are for highest-resolution shell.

Despite several co-crystallization and soaking experiments with substrate analogs mentioned in Table 2.2, we were unable to obtain a substrate bound structure of the enzyme. We propose that the substrate-binding site is located in the huge empty pocket above FMN. Also, substrate binding may involve conformational changes such that the residues on the top domain will close above FMN and substrate during catalysis. We

have recently observed such movement of a lid domain after binding of the substrate with riboflavin lyase (RcaE), which is discussed in the next section.

2.3.7 Reduction of sulfenic acid with CmoI (glutaredoxin)

Sulfenic acid **49** was demonstrated to be the product of CmoJ catalyzed reaction using several derivatization strategies. To prevent any damage due to the reactive nature of the sulfenic acid, this pathway encodes a glutaredoxin enzyme CmoI that might play a role in reduction of **49** to N-acetyl-L-cysteine (**38**).[44] As shown in Figure 2.18, the activity of CmoI was demonstrated by performing a coupled enzymatic reaction with CmoJ and derivatization of the product **38** using NBD-Cl (**60**).[45]

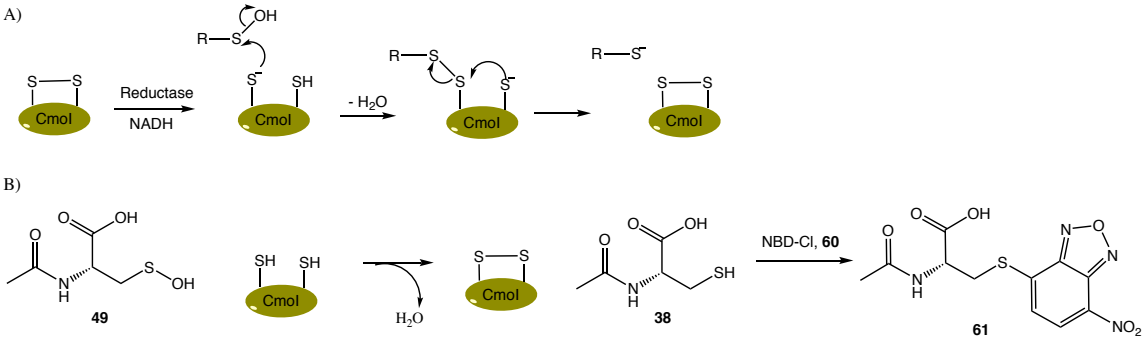


Figure 2.18: CmoI catalyzed reduction of sulfenic acid. A) The active site disulfide of CmoI can be reduced with a reductase *in vivo*. The reduced thiol will attack the sulfenic acid and form a crosslink. A second thiol in the active site can attack the disulfide to liberate the small molecule. B) CmoI catalyzed reduction of **49** during coupled reaction with CmoJ. Product **38** was derivatized with **60** which was confirmed by LC-MS. C) LC-MS analysis showing EIC $[M-H]^-$ = 325 Da demonstrating formation of **61** during CmoI-CmoJ coupled reaction. The inset shows MS of the derivatized reaction product **61** in negative mode.

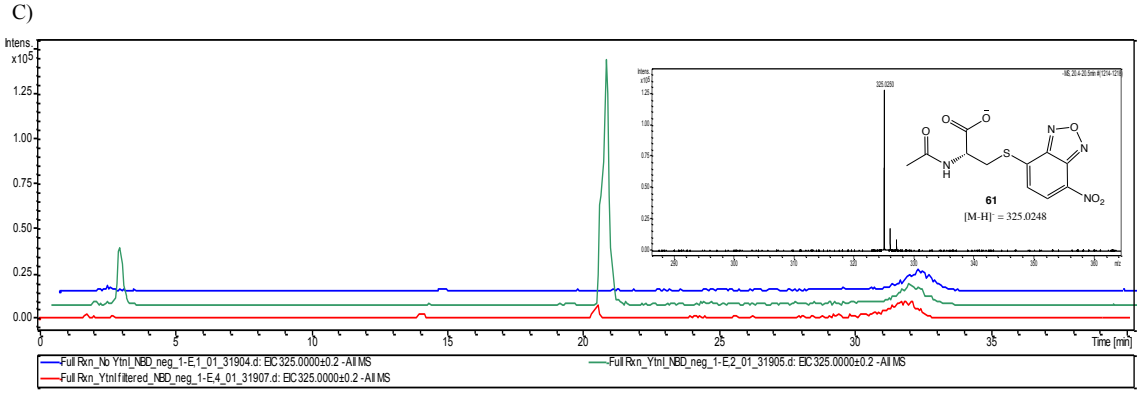


Figure 2.18 continued.

2.4 Conclusion

We have demonstrated the roles of all enzymes involved in this recently discovered L-cysteine (**8**) salvage pathway. Mechanistic insights were obtained for key reactions. As shown in Figure 2.19, acetylation by SnaA during the first step channels the right substrate to the rest of the pathway. Additionally, it may also prevent the toxic effects due to similarity of these non-canonical amino acids to L-cysteine (**8**) and L-methionine (**5**). Compound **29** is further oxidized to a stereospecific sulfoxide **67** by the flavin-monooxygenase CmoO. The second flavin-monooxygenase CmoJ, catalyzes the Pummerer-type rearrangement and generates sulfenic acid **49**. The highly reactive **49** is further reduced by CmoI and is deacetylated in the final step by SndA generating L-cysteine (**8**).

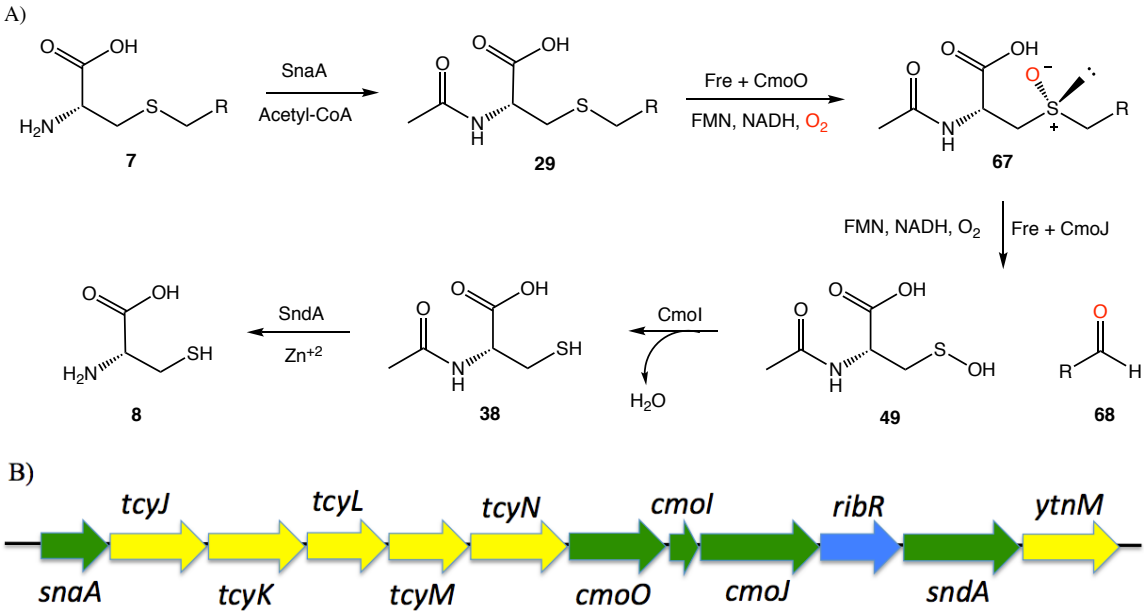


Figure 2.19: Role of the enzymes involved in cysteine salvage pathway. A) Substrates and products of enzymes in the pathway. B) Gene cluster involved in cysteine salvage pathway (enzymes are color coded as green).

Kinetic studies using various substrate analogs for the CmoJ catalyzed reaction showed enhanced activity with S-benzyl substituted cysteine. S-benzyl-L-cysteine was also found to be an efficient substrate for the SnaA catalyzed acetylation followed by the CmoO catalyzed sulfoxidation. This suggests that the S-benzyl substituted cysteines are preferred substrates of the enzymes in this pathway. Since this pathway is involved in salvage of L-cysteine (**8**) it is interesting to note that both the flavoenzymes CmoO and CmoJ do not contain any cysteine residues.

Previous studies on C-S lyases have predominantly focused on PLP dependent enzymes like alliinase[46] and EgtE in Ergothionine biosynthesis.[47] In Figure 2.1C we saw

examples of flavin mediated C-S lyases like the DszA involved in dibenzothiophene catabolism[48] and SsuD involved in alkanesulfonate metabolism.[49] Other reported examples include prenylcysteine lyase which is an FAD oxidase and doesn't generate sulfoxide nor requires NAD(P)H for its activity.[50] This study expands the repertoire of enzymes from flavin superfamily, which can catalyze C-S lyase activity by modulating the accessible redox states of sulfur.

The cysteine salvage pathway sheds light on a survival strategy evolved by microbes to obtain sulfur from organosulfur molecules produced by plants. Bioinformatic analysis has shown that this pathway is found in organisms that are closely associated with the roots of plants.[14] The pathway is upregulated in absence of readily available sulfur sources in the environment. The transporter and periplasmic proteins in the operon, play a role in the intake of S-substituted cysteines (**7**).[51] Recently, RibR was characterized as a dual functional enzyme demonstrating riboflavin kinase activity and regulating biosynthesis of flavins by binding FMN riboswitches.[52] This pathway uses various S-substituted cysteines (**7**) available in the environment as substrates and generates L-cysteine (**8**), which serves as a source of sulfur for any sulfur containing metabolite in those bacteria.

2.5 Experimental procedures

Materials: All chemicals were purchased from Sigma-Aldrich (now MilliporeSigma) unless specified otherwise. LB broth (Lennox formulation) was from EMD Millipore. Kanamycin was from Teknova and IPTG was obtained from Lab

scientific Inc. Chloramphenicol was from Fisher Scientific. HPLC and LC-MS solvents were purchased from EMD and were used without further purification. Histrap column (5 ml) was obtained from GE healthcare. Econo-Pack 10DG and Bio-spin 6 desalting columns were purchased from Bio-Rad Laboratories. Large cultures were grown and overexpressed in 2.5 L baffled ultra yield flasks from Thomson Instrument Company. NMR tubes with diameter of 3 mm and 5 mm were obtained from Wilmad-Labglass. D_2O, MeOD and d_6-DMSO were purchased from Cambridge Isotope Laboratories Inc. *MsrA* was gene synthesized and cloned in pTHT vector [modified pET28b with TEV protease cleavage site after N-terminal His-tag] by GenScript.

Overexpression and purification of enzymes: Genes of *snaA, cmoO, cmoI, cmoJ* and *sndA* were cloned in pTHT vector (derivative of pET28b vector with TEV protease cleavage site after N-terminal His-tag) while TEV protease was cloned in pET28b vector. The respective plasmids were transformed into BL21(DE3) competent cells of *E.coli* using electroporation technique. Starter cultures were grown overnight in LB media containing kanamycin (40 μg/ml). 30 ml of this culture was added to 3L LB media (2 x 1.5L flasks) with antibiotics and grown at 37 °C with shaking (180 rpm) till OD_{600}~0.6. The flasks were then incubated at 4 °C for ~2 hr without shaking. Cultures were then induced with 500 μM IPTG followed by incubation at 15 °C for ~14 hr with shaking at 180 rpm. The cells were harvested by centrifugation at 5,000 rpm for 20 min and stored in liquid nitrogen until further use. Typical yields were 9-10 gm of cell pellet (wet weight) from 3L cell culture.

For purification, the cell pellets were thawed and resuspended in 40-50 ml of lysis buffer (100 mM KPi, 150 mM NaCl, pH 7.5) at room temperature in the presence of lysozyme (6-8 mg). The suspension was stirred for ~1.5 hr on an ice-bath and further sonicated to lyse the cells. Cell debris was removed by centrifugation at 15,000 rpm for 20min and the lysate was filtered using 0.22 μm filters. The filtered lysate was loaded onto a His-trap column pre-equilibrated with lysis buffer. The column was washed with 10 column volumes of wash buffer (100 mM KPi, 20 mM imidazole, 150 mM NaCl, pH 7.5). The protein was then eluted from the his-trap column with elution buffer (100 mM KPi, 250 mM imidazole, 150 mM NaCl, pH 7.5). Protein containing fractions which were identified using Bradford reagent, were pooled together and concentrated using 15 ml 10kda filters. The concentrated protein was buffer exchanged to 50 mM KPi, 10% glycerol, pH 7.5 using the Econo-Pac 10DG desalting column. Aliquots of 100 μl desalted enzyme in eppendorf tubes were flash frozen in liquid nitrogen and stored in -80 °C until further use. Protein concentration was determined by measuring the absorbance at 280 nm (A_{280}) and utilizing the extinction coefficient calculated using ProtParam tool of the ExPASy proteomics server.

Co-expression of sndA with tig chaperone: Plasmid pTF16 encoding *tig* chaperone was obtained commercially (TaKaRa Bio Inc. chaperone plasmid set).[53] Competent cells of BL21(DE3) were transformed with pTF16 followed by plasmid containing *sndA*. Starter cultures were grown overnight in LB media containing kanamycin (40 μg/ml) and chloramphenicol (25 μg/ml). 30 ml of this culture was added to 3L LB media (2 x 1.5L flasks) containing 3 mg/ml L-arabinose and antibiotics.

The flasks were grown at 37 °C with shaking (180 rpm) till OD_{600} ~0.6 and further incubated at 4 °C for ~2 hr without shaking. Cultures were then induced with 500 µM IPTG followed by incubation at 15 °C for ~14 hr with shaking at 180 rpm. The cells were harvested by centrifugation at 5,000 rpm for 20 min and stored in liquid nitrogen until further use. Purification of enzyme using His-tag column was performed using the procedure mentioned in the previous page.

Purification of CmoJ for crystallographic studies: The cells were overexpressed and purified as described in the previous page. However, Tris buffer was used instead of KPi buffer throughout purification. After desalting CmoJ, purified TEV protease was added to it in the ratio TEV:CmoJ = 1:10 and incubated at 30 °C for 4 hr. The solution of proteins was passed through Ni-NTA column, which was pre equilibrated with (100 mM Tris, 150 mM NaCl, pH 7.5). Histag cleaved CmoJ was eluted from the column using 100 mM Tris, 20 mM imidazole, 150 mM NaCl, pH 7.5 and further desalted into 10 mM Tris, pH 7.5 buffer. The enzyme was either used directly for setting up crystallization trials or stored in 100 µl aliquots at -80 °C and used later.

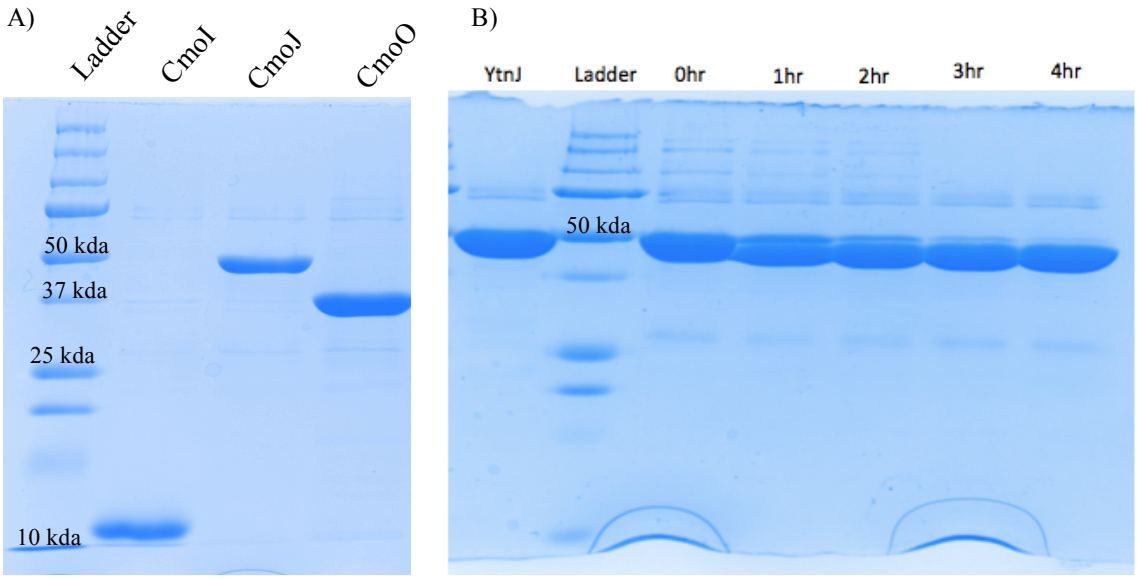


Figure 2.20: SDS-PAGE gel of purified proteins. A) Gel showing ladder (from Bio-rad), CmoI (10.5 kda), CmoJ (50 kda) and CmoO (37 kda) respectively. B) Time-course showing His-tag cleavage of CmoJ using TEV-protease (used for crystallographic study). His-tag is completely cleaved after ~4 hr.

HPLC parameters: An Agilent 1200 or 1260 HPLC equipped with a quaternary pump was used. The system included a diode array UV-Vis detector and chromatograms were detected using absorbance at 220 nm, 254 nm, 260 nm, 280 nm, 340 nm and 425 nm. Analysis was performed using either SUPELCOSIL LC-18-T (15 cm x 4.6 mm, 3 µm particles, Sigma) or ZORBAX Eclipse XDB-C18 column (15 cm x 4.6 mm, 5 µm particles, Agilent Technologies). Data was processed using ChemStation ver. B.04.01 SP1 (Agilent technologies).

***HPLC conditions**:*

A- Water

B- 100 mM Potassium phosphate buffer, pH 6.6

C- Methanol

***HPLC method**:* (Flow rate = 1 ml/min)

0 min – 100% B, 2 min – 10% A 90% B, 10 min – 45% A 15% B 40% C, 18 min – 25% A 15% B 60% C, 20 min - 25% A 15% B 60% C, 22 min - 100% B, 25 min - 100% B.

LC-MS parameters: LC-ESI-TOF-MS was performed using an Agilent 1260 HPLC system equipped with a binary pump and a 1200 series diode array detector followed by a MicroToF-Q II mass spectrometer (Bruker Daltonics) using an ESI source in negative mode. Analysis was performed on a LC-18-T column (15 cm x 3 mm, 3 μm particles, Supelco). Data was processed using DataAnalysis 4.0 SP1 (Bruker Daltonics).

***LC conditions**:*

A- 5 mM Ammonium acetate buffer, pH 6.6

B- 75% Methanol and 25% Water.

LC method: (Flow rate = 0.4 ml/min)

0 min – 100% A, 7 min – 100% A, 10 min – 80% A 20% B, 27 min – 100% B, 29 min – 100% B, 30 min – 100% A, 40 min – 100% A.

Reconstitution of SnaA catalyzed acetyl transferase: A reaction mixture (1 ml) consisting of 1 μM SnaA, 500 μM acetyl CoA (**33**), 500 μM DTNB (**35**, 100 mM stock

in DMSO) in 50 mM KPi buffer, pH 7.5 at 25 °C was added in a quartz cuvette. This mixture was used as a blank for UV-Vis spectrophotometer. The reaction was initiated by adding substrate/analog and the Absorbance was recorded at 412 nm for 10 min.

Reconstitution of SndA catalyzed deacetylation: A 100 µl aliquot of 75 µM purified SndA was incubated with 1.5 mM 8-hydroxy-5-quinolinesulfonic acid (divalent metal chelator) for few hrs. Protein was desalted using Bio-spin 6 desalting columns and enzymatic reaction was setup with various metals like Zn^{+2}, Mg^{+2}, Fe^{+2}, Fe^{+3}, Ni^{+2}, Mn^{+2}, Co^{+2}. A reaction mixture (100 µl) consisting of 10 µM SndA, 1 mM N-acetyl-L-cysteine (**38**), 1 mM TCEP, 1 mM metal solution in 50 mM KPi buffer, pH 7.5 at 25 °C was incubated for 2 hr. Protein was removed from sample using a 10 kda filter. A 50 mM Bromobimane (**40**) solution in DMSO was prepared and 4 µl was added to each sample. Samples were analyzed on LC-MS by injecting 80 µl of the reaction mixture.

Reconstitution of CmoO catalyzed sulfoxidation: Samples for NMR analysis were prepared by separately desalting Flavin reductase (Fre) and CmoO using Bio-spin 6 desalting columns to remove glycerol from protein solution. A reaction mixture (400 µl) consisting of 1 µM Fre, 40 µM CmoO, 250 µM FMN, 10 mM NADH, 5 mM N-acetyl-S-methyl-L-cysteine (**38**) in 50 mM KPi buffer, pH 7.5 at 25 °C was incubated for 30 min. 100 µl of D_2O-KPi buffer, pD 7.1 was added to enzymatic reaction and the samples were transferred to NMR tubes. NMR was recorded using Bruker Avance III 400 MHz instrument. To identify the stereoisomer generated during CmoO catalyzed reaction, the

'Full reaction' sample was spiked with synthetic diastereomers of N-acetyl-S-methyl-L-cysteine sulfoxide (**43**) and NMR was recorded.

Samples prepared for HPLC / LC-MS analysis had the following composition. A reaction mixture (100 µl) consisting of 150 nM Fre, 10 µM CmoO, 25 µM FMN, 1 mM NADH, 500 µM N-acetyl-S-benzyl-L-cysteine in 50 mM KPi buffer, pH 7.5 at 25 °C was incubated for 30 min. Protein was removed using 10 kda filter and 40-80 µl of sample was injected in HPLC / LC-MS.

Determining stereochemistry of sulfoxide by MsrA: Samples for NMR analysis were prepared by desalting MsrA using Bio-spin 6 desalting columns to remove glycerol. NMR was recorded of 100 µl of D_2O-KPi buffer added to a reaction mixture (450 µl) consisting of 20 µM MsrA, 3 mM DTT and 10 mM N-acetyl-S-methyl-L-cysteine sulfoxide (**43**) in 50mM KPi buffer, pH 7.5 at 25 °C which was incubated for 30min. Additional 3 mM DTT was added to let the reaction go to completion and NMR of the sample was recorded again.

Reconstitution of the flavin monooxygenase CmoJ: Samples for NMR analysis were prepared by desalting Flavin reductase (Fre) and CmoJ using Bio-spin 6 desalting columns to remove glycerol. A reaction mixture (400 µl) consisting of 1 µM Fre, 40 µM CmoJ, 250 µM FMN, 10 mM NADH, 5 mM N-acetyl-S-methyl-L-cysteine sulfoxide (**43**) or N-acetyl-S-benzyl-L-cysteine sulfoxide (**46**) in 50 mM KPi buffer, pH 7.5 at 25 °C was incubated for 2-18 hr. 100 µl of D_2O-KPi buffer, pD 7.1 was added to each sample and NMR spectra was recorded.

Identity of benzaldehyde (**50**) and N-acetyl cysteine (**38**, after addition of TCEP) as products of CmoJ catalyzed reaction was established by recording NMR spectra of the product standards individually followed by spiking the 'Full reaction' sample with these compounds.

Samples prepared for HPLC / LC-MS analysis had the following composition. A reaction mixture (100 µl) consisting of 150 nM Fre, 10 µM CmoJ, 25 µM FMN, 1 mM NADH, 500 µM N-acetyl-S-benzyl-L-cysteine sulfoxide (**46**) in 50 mM KPi buffer, pH 7.5 at 25 °C was incubated for 30 min. Protein was removed using 10 kda filter and 40-80 µl of sample was injected in HPLC / LC-MS. Following variations were done to the enzymatic reaction set-up to derivatize the sulfenic acid generated during CmoJ catalyzed reaction: 2 mM dimedone (**56**) was included at the start of the reaction to form adduct **57**, 2 mM 4-mercaptopyridine (**58**) was included the start of the reaction to form adduct **59** while 2 mM TCEP was included at the start of the reaction and 2 mM NBD-Cl (**60**) dissolved in DMSO was added after removing the enzymes using 10 kda filter to generate adduct **61**.

For the $^{18}O_2$ experiment, all buffers and reagents were transferred into the anaerobic chamber ($\leq$ 5 ppm O_2, COY Laboratories). Fre and CmoJ were desalted using Bio-spin 6 desalting columns into anaerobic buffers. Stock solutions of substrates and cofactors were also prepared in anaerobic 50 mM KPi buffer, pH 7.5. A reaction mixture (100 µl) consisting of 150 nM Fre, 10 µM CmoJ, 25 µM FMN, 1 mM NADH, 500 µM N-acetyl-S-(p-hydroxybenzyl)-L-cysteine sulfoxide (**62**) were setup in an eppendorf tube in two sets and sealed. The eppendorf tubes were removed from anaerobic chamber and

exposed with a ballon of $^{16}O_2$ for one set and $^{18}O_2$ for the second one. After 10 min, the reaction was quenched and extracted with 2 x 100 µl of ethyl acetate. The ethyl acetate fraction was concentrated and treated with 2ul of BSTFA to derivatize the hydroxyl group of p-hydroxybenzaldehyde (**63**) at 60 °C for 15min.[54] The samples were further analyzed using GC-MS.

Synthesis of substrates and products: Synthesis involved formation of S-substituted-L-cysteine (**7**), followed by acetylation or sulfoxidation (sequence of acetylation and sulfoxidation can be used interchangeably). All steps were performed using procedures reported in literature. L-cysteine (**8**) S-methyl-L-cysteine (**26**), S-carboxymethyl-L-cysteine (**28**) and S-benzyl-L-cysteine (**44**) were obtained commercially from Sigma-Aldrich. The general reaction scheme is shown in Figure 2.21 and representative ^{1}H-NMR are shown in Figures 2.22 and 2.23.

Step 1: Preparation of aromatic S-substituted-L-cysteine.[55] The reaction was setup under Argon atmosphere. In a 25 ml round bottom flask, dissolve 0.2 mmoles of L-cysteine.HCl in 10 ml of 2N NaOH and add 0.2 mmoles of 4-hydroxybenzyl chloride (precursor for **62**) or 4-methoxybenzyl chloride (precursor for **66**). To this mixture, add 1 ml of dimethoxyethane. Reaction was allowed to progress for 30 min and was monitored using NMR. Once the reaction had gone to completion, the product was extracted using ethyl acetate and bulk of the solvent was evaporated using rotovap. Hexane was added dropwise to the ethyl acetate fraction to obtain the product as precipitate. The product was filtered using Buchner funnel and used for the next step.

Step 2: Acetylation of S-substituted-L-cysteine (**7**).[56] In a round bottom flask add 10 ml acetic acid. Dissolve 0.2 mmoles of a particular S-substituted-L-cysteine (**7**) and add 0.22 mmoles of acetic anhydride. Heat the reaction mixture to 60 °C for 30 min. Rotovap the reaction mixture to remove excess acetic acid and acetic anhydride.

Step 3: Sulfoxidation of the N-acetyl-S-substituted cysteine (**29**).[57] In a round bottom flask add 10 ml acetic acid to dissolve 0.2 mmoles of a particular N-acetyl-S-substituted-L-cysteine (**29**). Add 100 µl H_2O_2 solution (33% w/v) and stir at 60 °C for 30 min. The progress of reaction was monitored using ^{1}H-NMR.

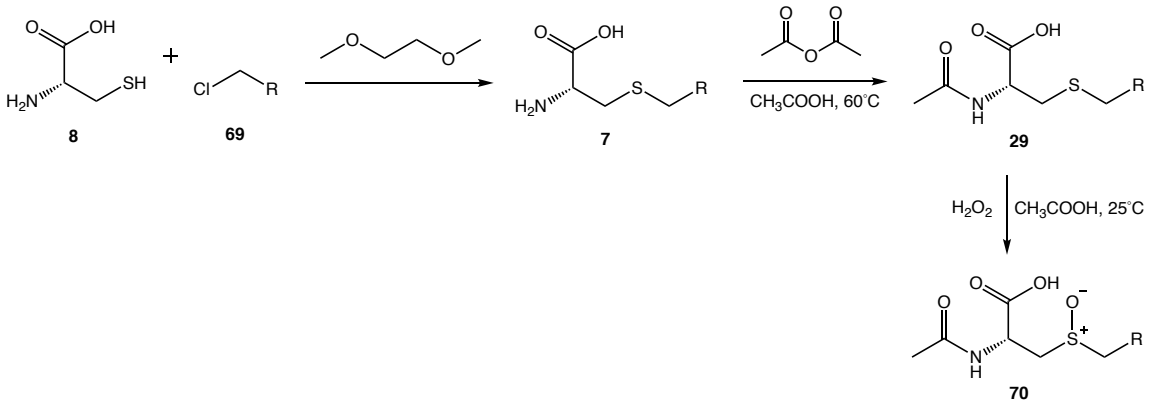


Figure 2.21: Synthetic scheme for substrates and products.

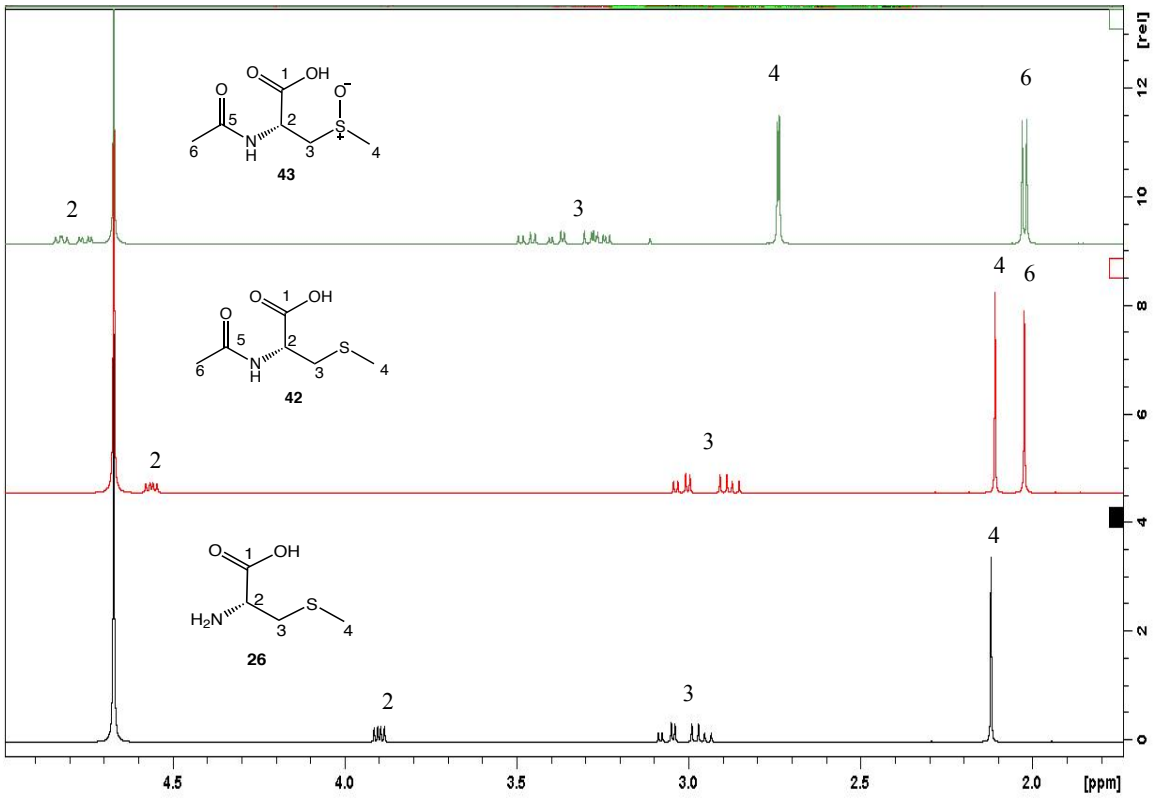


Figure 2.22: ^{1}H-NMR spectra of S-methyl-L-cysteine (**26**), N-acetyl-S-methyl-L-cysteine (**42**) and N-acetyl-S-methyl-L-cysteine sulfoxide (**43**) recorded in D_2O.

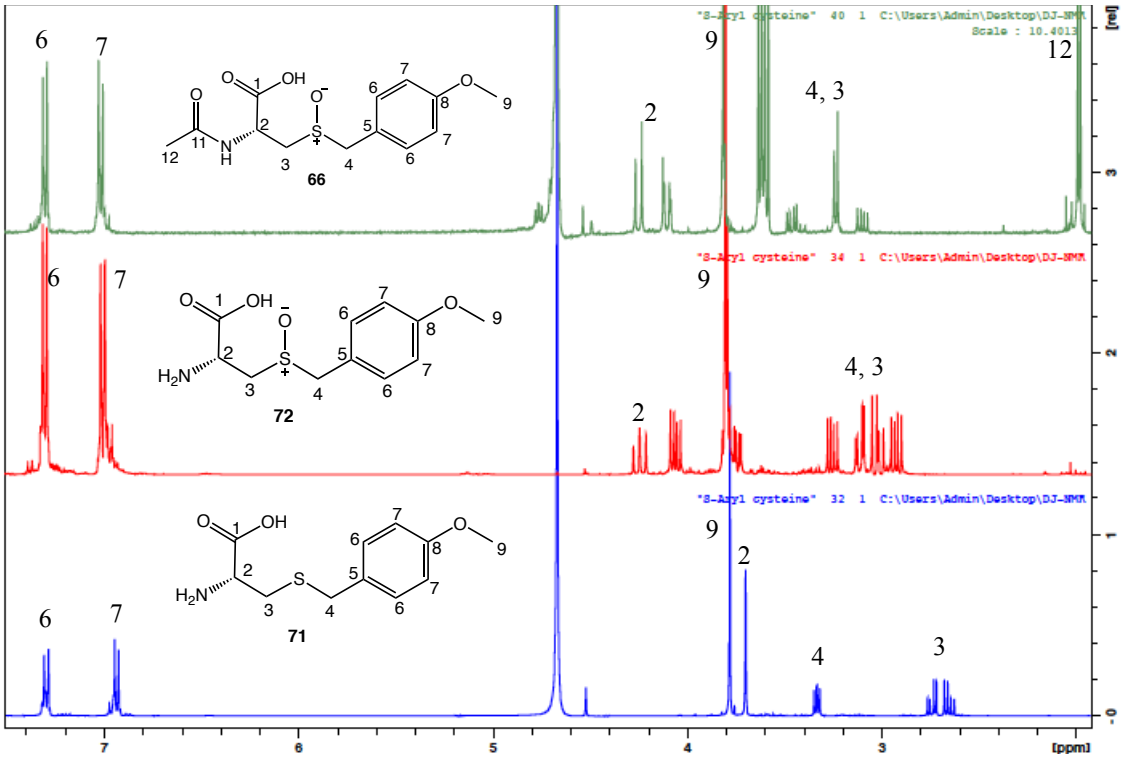


Figure 2.23: ^{1}H-NMR spectra of S-(p-methoxybenzyl)-L-cysteine (**26**), S-(p-methoxybenzyl)-L-cysteine sulfoxide (**42**) and N-acetyl-S-(p-methoxybenzyl)-L-cysteine (**43**) recorded in D_2O.

Synthesis of the adduct 59 between N-acetyl-L-cysteine (38) and 4-mercaptopyridine: In a round bottom flask, under Argon atmosphere, add 10 ml THF (tetrahydrofuran). Dissolve 0.2 mmoles of N-acetyl-L-cysteine (**38**) and 0.22 mmoles of aldrithiol-4 (**73**) and stir the reaction mixture at 25 °C for 24 hr. Remove the solvent using rotovap and dissolve the residue in KPi buffer. Adduct **59** was purified using HPLC and was further used to record NMR, co-injection and prepare calibration curve for kinetic experiments.

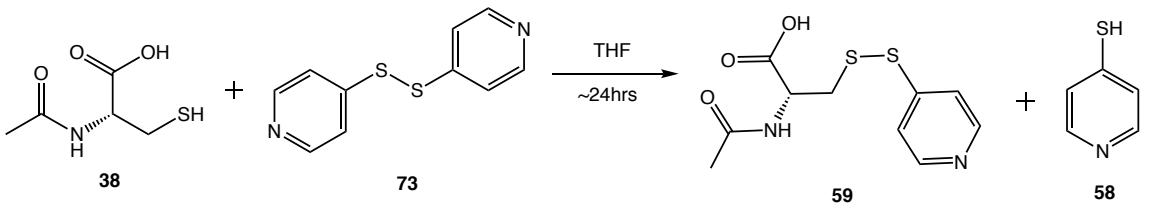


Figure 2.24: Synthetic scheme for adduct **59**.

Kinetic studies on CmoJ catalyzed reaction: A reaction mixture (100 µl) consisting of 150 nM Fre, 100-2500 nM CmoJ, 20 µM FMN, 1-2000 µM substrate and 1 mM 4-mercaptopyridine (**58**) in 50 mM KPi buffer, pH 7.5 at 25 °C was prepared. The enzymatic reaction was initiated by addition of 1-2 mM NADH. The reaction was quenched at suitable time-points (30-600 sec) using 1:1 ratio of enzymatic reaction and 8 M guanidine hydrochloride. The enzyme was removed by filtration and analysis of samples was performed using HPLC. The amount of adduct **59** was calculated by determining the area under the signal at 254 nm. The experiments were performed in triplicates for each substrate analog.

Crystal structure of CmoJ: Crystallization trials were setup using commercial kits and hanging drop diffusion method. Enzyme concentration was kept close to ~15 mg/ml. Few hits were obtained at 4 °C and were optimized further optimized to yield better quality crystals.

Apo CmoJ: 2.5 µl of reservoir solution was added to 2.5 µl of purified CmoJ on the coverslip and sealed on top of a reservoir containing 50 mM HEPES pH 7.5, 200

mM KCl and 29%-33% of pentaerythritol propoxylate (5/4 PO/OH) at 4 °C. Diamond shaped crystals were obtained after one week which grew larger for ~15 days.

CmoJ with FMN: The apo crystals of CmoJ were soaked overnight in 4 mM FMN dissolved in crystallization buffer.

Glycerol was used as a cryoprotectant while preparing samples for X-ray diffraction.

Collection of X-ray diffraction data: All X-ray diffraction data was collected at the Advanced Photon Source (APS), Argonne National laboratory. Crystals of the apo CmoJ were diffracted at 100K, at the 23-ID-D beamline using the PILATUS3 6M detector and a wavelength of λ= 1.033 Å. Crystals of FMN bound CmoJ were diffracted at 100K, at the 19-ID beamline using the DECTRIS PILATUS 2M detector and a wavelength of λ= 0.925 Å. Diffraction data was collected for 240-360 frames and was recorded in a shutterless mode at the rate of one degree/sec. Datasets were integrated and scaled using iMosflm for the apo CmoJ and HKL2000 for the FMN bound CmoJ.

Structure determination and refinement: The structure of CmoJ bound with FMN was determined by molecular replacement using previously deposited structure of the same enzyme PDB id 1YW1 as the starting model. The structures were refined iteratively using programs like PHENIX and COOT. The model thus generated was used as a starting model for CmoJ (apo form) and CmoJ bound with FMN. The final refinement statistics are summarized in Table 2.3.

3. STUDIES ON RIBOFLAVIN CATABOLIC PATHWAY AND STRUCTURAL ENZYMOLOGY OF RIBOFLAVIN LYASE

3.1 Catabolism of cofactors[58]

Genes involved in the biosynthesis of most cofactors were identified in last few decades and the mechanistic studies in maximum cases are at advanced stages. These studies unraveled a large volume of novel organic chemistry that was unprecedented in the biological realm.[59] However, despite early reports on the catabolism of several cofactors, gene clusters of only heme, pyridoxal phosphate and NAD^+ have been reported and studied. Bacteria involved in degradation of thiamin, folate, riboflavin and biotin were once isolated but have now been lost. The current understanding of cofactor catabolism is still in its infancy and has a potential to uncover remarkable series of unconventional transformations.

3.2 Gene cluster involved in catabolism of riboflavin (1)

In 1944, J. W. Foster identified a strain that could degrade riboflavin (**1**) and named the organism *Pseudomonas riboflavina* (later renamed as *Devosia riboflavina* based on 16S rRNA sequence).[15, 60] Further studies with crude cell lysates suggested ribitol and lumichrome as products from crude cell extracts. However, the enzymes responsible for this transformation remained elusive. In the meantime, additional riboflavin degradation pathways were discovered but these strains were lost. In last few decades, no progress has been reported on genes and pathways involved in riboflavin catabolism.

3.2.1 Strain *Microbacterium maritypicum* G10 catabolizes riboflavin (1) to lumichrome (11)[19]

To identify strains that could catabolize riboflavin, dust samples were collected from a DSM riboflavin production plant in Switzerland by Dr. Benjamin Philmus (former postdoc in Begley lab). These samples were screened in M9 minimal media for growth in presence of riboflavin (**1**) as a sole carbon source. In few samples, bleaching of the yellow chromophore of riboflavin was observed. One of the strains that showed a robust growth on riboflavin was chosen for further studies. The 16S rRNA sequence identified this strain as *Microbacterium maritypicum.*

Next, I investigated whether riboflavin could be utilized as a source of carbon as well as nitrogen by *M. maritypicum*. When NH_4Cl was excluded from the M9 media supplemented with riboflavin, no growth of this strain was observed. This suggested that the catabolic pathway is unable to extract the nitrogen's present in the 3-membered isoalloxazine ring of riboflavin and utilize it for growth. As shown in Figure 3.1, further experiments on varying riboflavin concentrations in M9 minimal media demonstrated that higher concentration of riboflavin in growth media led to higher OD_{600} values.

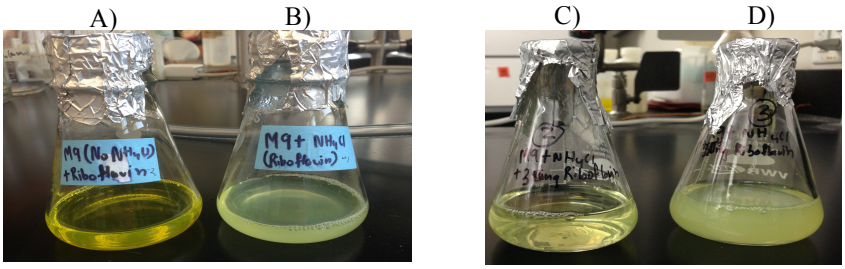


Figure 3.1: Growth of *M. maritypicum* in M9 minimal media supplemented with riboflavin (**1**). A) M9 media (no NH_4Cl) supplemented with riboflavin, no growth observed. B) M9 media (with NH_4Cl) supplemented with riboflavin, growth observed. C) M9 media + NH_4Cl + 3mg riboflavin, bleaching of chromophore and weak growth. D) M9 media + NH_4Cl + 5mg riboflavin, bleaching of chromophore and robust growth.

To identify the products generated from degradation of riboflavin (**1**), the spent media was analyzed for metabolites. HPLC analysis of the media showed time-dependent consumption of **1** and appearance of a new signal, which comigrated with lumichrome (**11**, Figure 3.1A). Analysis of sample after 21 hr showed complete consumption of **1**, a small peak for **11** and no additional signals were observed on the HPLC. This suggested that either lumichrome (**11**) is completely consumed or it precipitates from the growth media. To test for presence of lumichrome, the spent media was divided in two fractions. One fraction was filtered (as previously done for HPLC samples) while the other was used directly. Both fractions were treated with 1:1 ratio of spent media and 1 N KOH solution (KOH increases solubility of lumichrome to 50 mg/ml).[61] UV-Vis spectra of these samples were recorded. As shown in Figure 3.1B, lumichrome was observed in sample that was not filtered while the sample which was filtered, showed trace amounts of lumichrome (**11**). Further quantitation of **11** in unfiltered sample accounted for ~96% of riboflavin (**1**) initially added in the media.

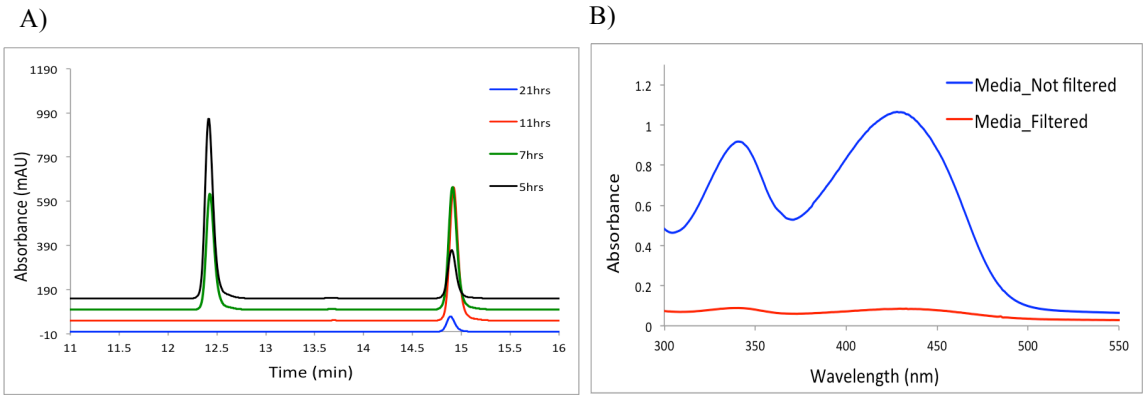


Figure 3.2: Degradation of riboflavin (**1**) to lumichrome (**11**) by *M. maritypicum*. A) HPLC analysis at 340 nm showing time dependent formation of lumichrome from riboflavin. UV-Vis analysis of culture medium treated with 1N KOH after ~30 hr of growth.

3.2.2 Riboflavin catabolic gene cluster in *Microbacterium maritypicum* G10

The growth studies of *M. maritypicum* in media supplemented with riboflavin (**1**) suggested that this pathway involves cleavage of the ribose from the isoalloxazine ring and lumichrome (**11**) is not further utilized as a source of carbon or nitrogen (Figure 3.3A). To identify the genes involved in this transformation, Dr. Hui Xu (former postdoc in Begley lab) prepared cosmid library of *M. maritypicum* genome in *Streptomyces lividans* as heterologous host and screened for riboflavin catabolism. Two different cosmid hits were identified during the screening process. Further experiments identified a 8.7kb fragment, which showed degradation of riboflavin to lumichrome (Figure 3.3B). The genes were annotated based on the biochemical characterization of enzymes in the pathway performed by Dr. Yindrila Chakrabarty (former graduate student in Begley lab).

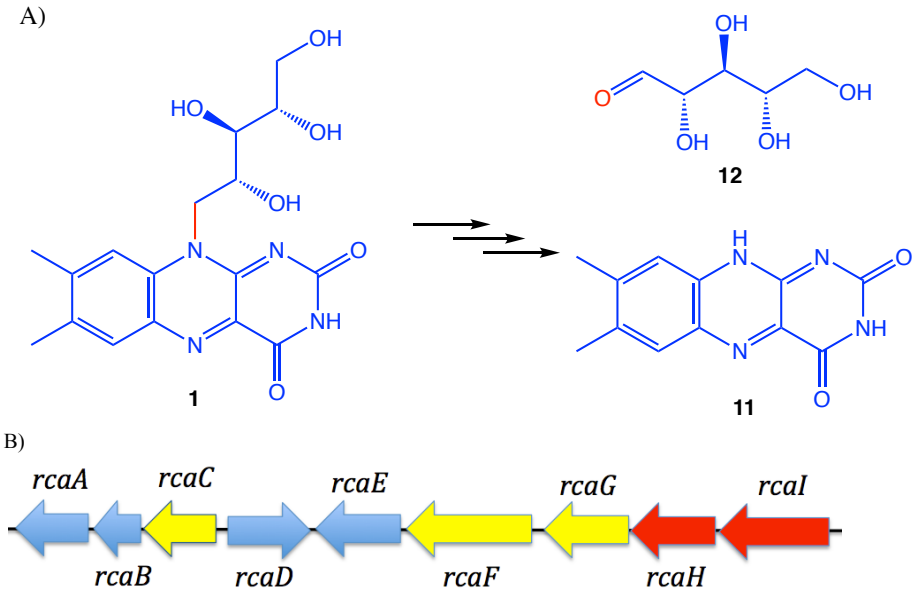


Figure 3.3: Catabolism of riboflavin (**1**) in *M. maritypicum.* A) Riboflavin is catabolized to ribose (**12**, used for growth) and lumichrome (**11**, discarded in media). B) Genes in the 8.7kb fragment that includes the cluster involved in the above mentioned transformation.

Table 3.1: Annotation of genes involved in riboflavin (**1**) catabolism based on biochemical characterization.

Sr. No.	gene	Function
1.	*rcaA*	Flavin kinase
2.	*rcaB*	Flavin reductase
3.	*rcaC*	PadR family transcriptional regulator
4.	*rcaD*	Ribokinase
5.	*rcaE*	Riboflavin lyase
6.	*rcaF*	Transporter
7.	*rcaG*	Transporter
8.	*rcaH*	Unknown function
9.	*rcaI*	Unknown function

3.2.3 Role of enzymes in the riboflavin catabolic gene cluster

The reactions catalyzed by the enzymes mentioned in Table 3.1 are highlighted in Figure 3.4. RcaC regulates the expression of the gene cluster in the environment. RcaF and RcaG are involved in transportation of riboflavin (**1**) into the cell. RcaA phosphorylates **1** to FMN (**2**) in presence of ATP and RcaB reduces this FMN (**2**) in presence of NADH to form $FMNH_2$ (**74**). RcaE utilizes this $FMNH_2$ (**74**) as a cofactor and acts on substrate riboflavin (**1**) in presence of oxygen to form lumichrome (**11**) and ribose (**12**). RcaD catalyzes phosphorylation of ribose (**12**) to form ribose-5-phosphate (**75**), which channeled into the central metabolism of an organism. Lumichrome (**11**) is released in the environment where it has been previously shown to be involved in quorum sensing. In summary, this pathway cannibalizes the ribose side chain of riboflavin (**1**), which serves as a carbon source for the growth and survival of bacteria.

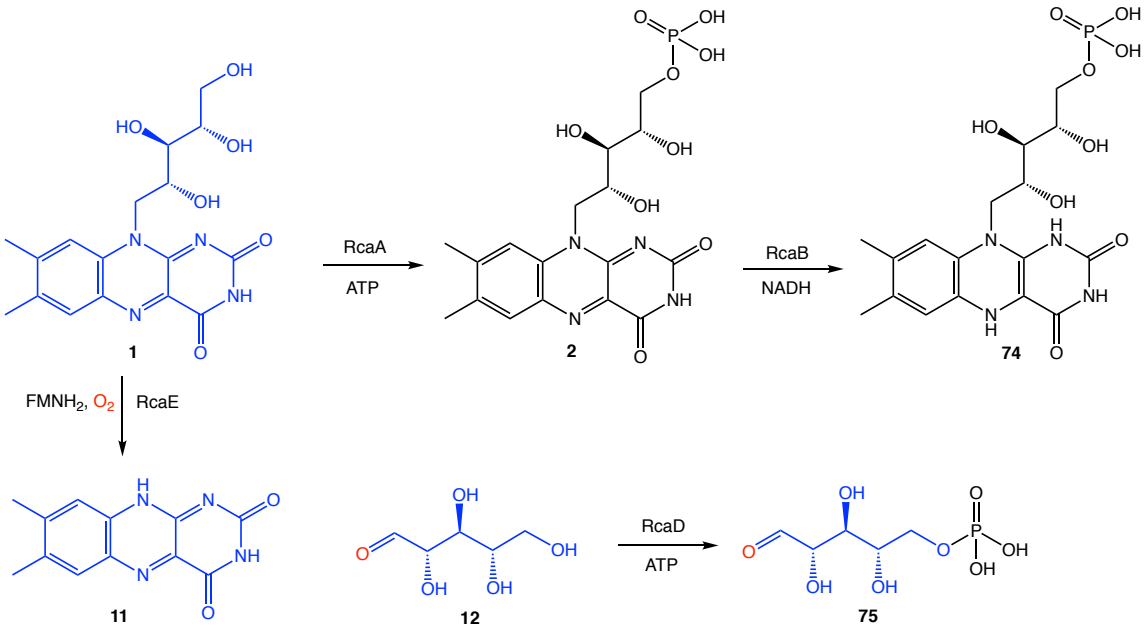


Figure 3.4: Role of all enzymes in the riboflavin catabolic gene cluster identified in *M. maritypicum*. RcaE cleaves the ribose from the isoalloxazine in a process involving phosphorylation (RcaA) and reduction (RcaB). Ribose (**12**) is further phosphorylated to ribose-5-phosphate (**75**, RcaD), which is used to sustain growth of the organism.

The fundamental reaction between reduced flavins and O_2 has been extensively studied since 1970s.[62] It has been primarily observed that $FMNH_2$ (**74**) initially transfers a single electron to O_2 and generates a radical pair between flavin semiquinone (**76**) and superoxide radical (**9**). In some enzymes like oxidases, a second electron transfer event generates H_2O_2 and oxidized flavin, while in monooxygenases, the radical pair collapses to form C_{4a}-flavin(hydro)peroxide (**77**). However, mechanistic studies on riboflavin lyase (RcaE) demonstrated that this enzyme utilizes the superoxide radical (**9**) generated using $FMNH_2$ (**74**) and O_2 to abstract a H-atom from the C_1' of riboflavin (**1**) as shown in Figure 3.5B. RcaE is the first example of flavin dependent enzyme where it has been

demonstrated biochemically that it catalyzes H-atom abstraction from the substrate using superoxide radical (**9**).

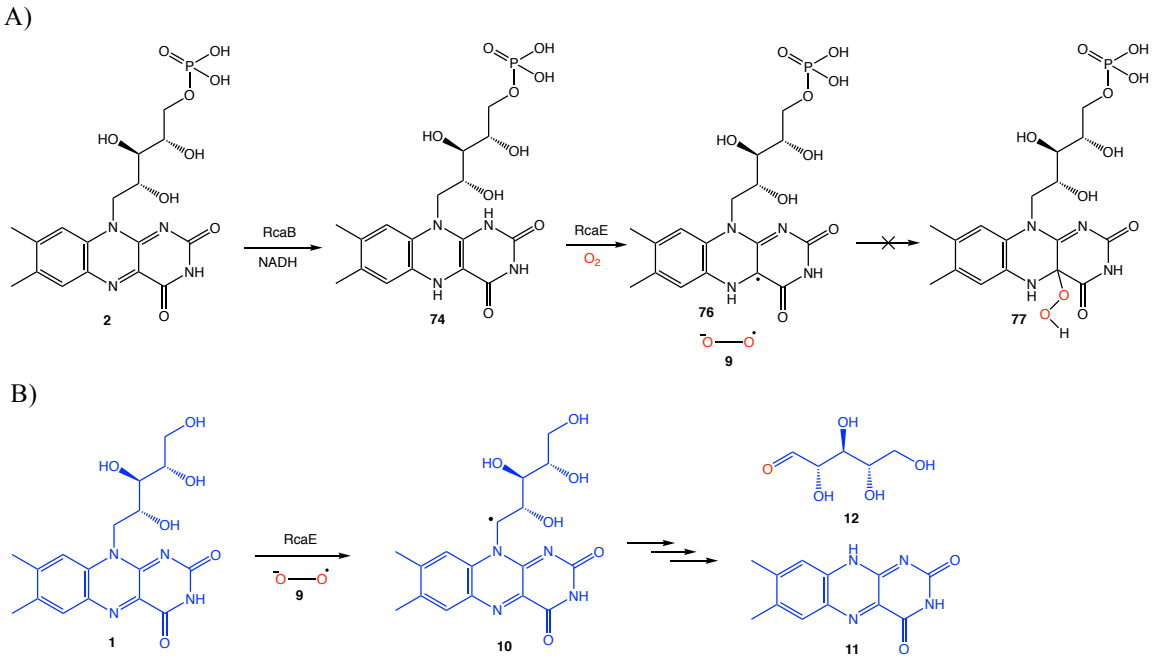


Figure 3.5: Superoxide radical mediated H-atom abstraction from C_1' of riboflavin during the RcaE catalyzed reaction. A) Reduction of FMN (**2**) generates $FMNH_2$ (**74**) which in presence of O_2 forms flavin semiquinone radical (**76**) and superoxide radical (**9**) pair. This step is common between other flavin enzymes and RcaE. However, in all other cases except RcaE, the radical pair collapses to form C_{4a}-flavin(hydro)peroxide (**77**). B) In RcaE, the superoxide radical (**9**) abstracts H-atom from the C_1' of riboflavin generating **10** which is subsequently converted to lumichrome (**11**) and ribose (**12**).

3.3 Crystal structure of Riboflavin lyase (RcaE)

To obtain further mechanistic insights into this superoxide radical (**9**) mediated transformation and to learn about the unique features of RcaE active site, we turned our attention to obtain the crystal structure of riboflavin lyase (RcaE). The crystals of apo RcaE were obtained at room temperature while the cofactor FMN was co-crystallized with RcaE at 4 °C as a dimer. Several attempts to obtain riboflavin (**1**) bound in the active site using co-crystallization and soaking methods were unsuccessful. One of the hypotheses for the failure to obtain **1** bound in the active site of RcaE was the low solubility (200 μM) of riboflavin (**1**) in buffer solutions. We developed several strategies to overcome the limitation of riboflavin solubility. One approach involved preparing riboflavin *in situ* from FMN (**2**) by adding a phosphatase. With this methodology, we were able to obtain ~2.5 mM RBF in buffer solution. Co-crystallization setup using this strategy yielded crystals with RBF (**1**) and FMN (**1**) bound in the active site of RcaE. The structures of all three forms of riboflavin lyase (RcaE) are shown in Figure 3.6.

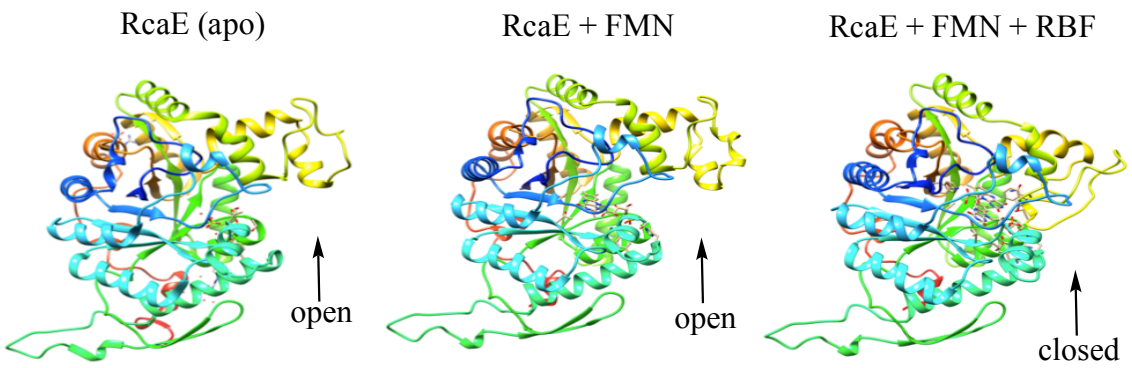


Figure 3.6: Structure of RcaE in its apo form (1.90 Å), FMN bound to RcaE (1.90 Å), and FMN + RBF bound to RcaE (1.75 Å).

As shown in Figure 3.6, structures of RcaE in the apo form and FMN bound form are quite identical. The active site in both cases is wide open and exposed to solvent. However after binding of the substrate RBF (**1**), the residues between Gln 306 to Pro 368 undergo movement and close on top of FMN + RBF thereby shielding it from bulk solvent. As shown in Figure 3.7, the electron density map of FMN + RBF bound RcaE illustrates that the phosphate group in FMN has migrated from C_5' to C_3' due to acidic crystallization buffer conditions (pH 6.0). The modified FMN was named ligand 9WY.

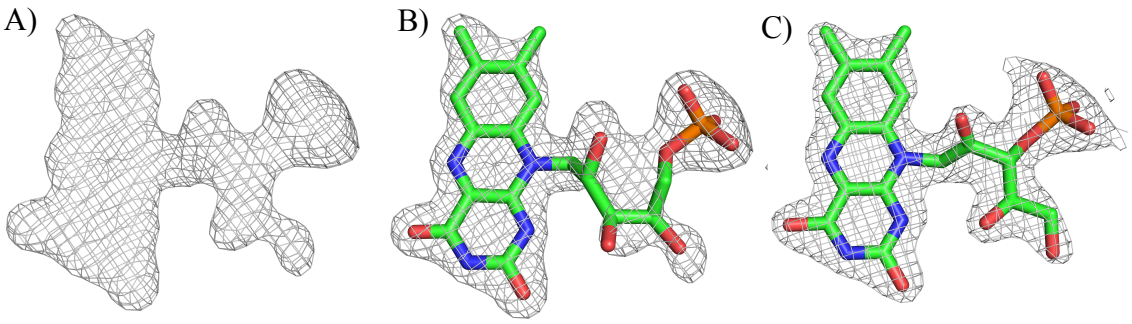


Figure 3.7: Modeling the cofactor ligand in the active site electron density of RcaE. A) $2F_O$-F_C map of the cofactor in active site of RcaE contoured at 1.0σ. B) Model of FMN in the $2F_O$-F_C map does not fit well. C) Model of 9WY fits accurately to the $2F_O$-F_C map.

Table 3.2: Data collection and refinement statistics for RcaE.

	RcaE (Apo form)	RcaE + FMN	RcaE + 9WY + RBF
Data collection			
Space group	$P2_12_12_1$	$P2_12_12_1$	$P2_12_12_1$
Cell dimensions			
a, *b*, *c* (Å)	74.97, 109.49, 126.42	76.25, 113.47, 117.81	74.97, 108.80, 125.08
α, β, γ (°)	90, 90, 90	90, 90, 90	90, 90, 90
Resolution (Å)	64.48-1.90 (1.96-1.90)*	64.01-1.90 (1.96-1.90)*	64.30-1.75 (1.81-1.75)*
R_{merge}	0.229 (2.965)	0.147 (1.758)	0.114 (3.591)
$I / \sigma I$	7.3 (1.3)	13.9 (2.7)	9.0 (0.8)
Completeness (%)	100 (99.92)	99.2 (99.93)	99.6 (99.37)
Redundancy	7.7 (7.9)	18.6 (18.6)	8.8 (8.8)
Refinement			
Resolution (Å)	1.90	1.90	1.75
No. reflections	82558	81065	102980
R_{work} / R_{free}	0.171/0.209	0.155/0.193	0.177/0.206
No. atoms			
Protein	7459	7375	7437
Ligand/ion	35	62	116
Water	490	377	379
B-factors			
Protein	32.51	47.73	47.51
Ligand/ion	56.73	60.56	47.93
Water	36.97	44.67	46.90
R.m.s. deviations			
Bond lengths (Å)	0.006	0.006	0.006
Bond angles (°)	0.845	0.840	0.853

*Values in parentheses are for highest-resolution shell.

The active site pocket of RcaE is hydrophobic and the isoalloxazine rings of the two flavins are stacked within Van der waals distance as shown in Figure 3.8A. We did not observe any electron density for water or O_2 near the reaction center i.e. near C_{4a} or N_5 of 9WY or C_1' of RBF. A molecule of O_2 is modeled close to this reactive center in the active site of RcaE (Figure 3.8B). During RcaE catalyzed reaction, a proton can be

transferred from N_5 of $FMNH_2$ (**74**) to the oxygen simultaneously while transferring an electron thus generating hydroperoxy radical (H-O-O•). This radical is optimally located to abstract a H-atom from the C_1' position of the substrate riboflavin (**1**). Thus the unique superoxide mediated transformation in RcaE occurs due to alignment of FMN (**2**) and RBF (**1**) such that the formation of C_{4a}-flavin(hydro)peroxide is sterically hindered which prevents biradical coupling between flavin semiquinone (**76**) and superoxide radical (**9**) species. Additionally, the C_1' position of riboflavin is optimally oriented for H-atom abstraction by the hydroperoxy radical (H-O-O•).

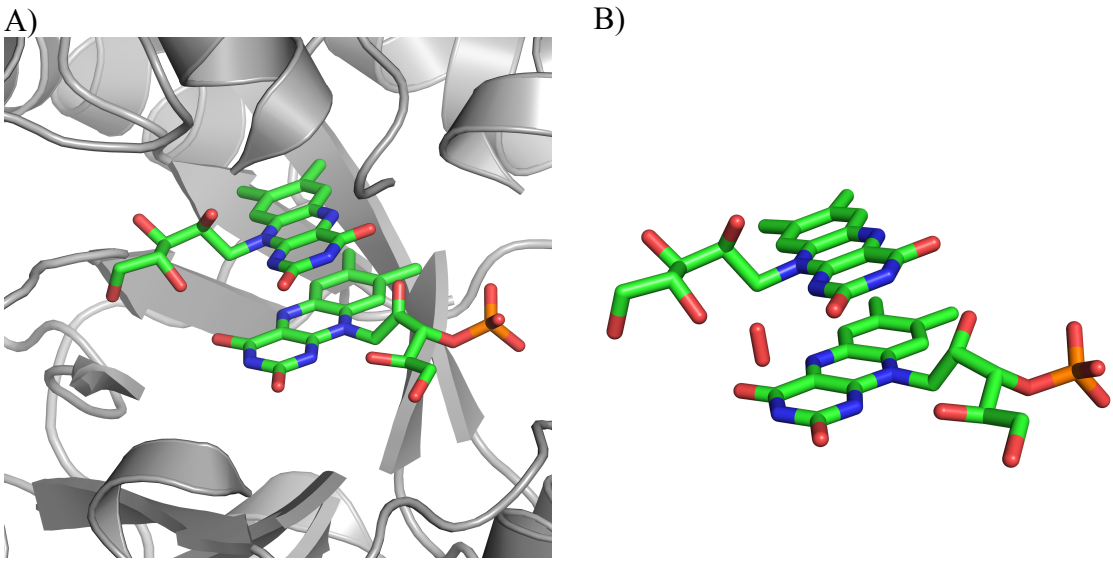


Figure 3.8: Orientations of FMN and RBF in the active site of RcaE. A) Modified FMN and RBF are stacked within Van der waals distance in RcaE active site. B) Oxygen modeled close to the reaction center of FMN (modified) and RBF in the active site of RcaE.

3.4 Conclusion

We have identified a novel strain *Microbacterium maritypicum* G10 that is capable of catabolizing riboflavin (**1**, Vitamin B_2). The catabolic pathway involves cleavage of the ribose from the isoalloxazine ring of **1**. The ribose (**12**) is used to sustain cell growth and the lumichrome (**11**) precipitates from the growth medium. The genes involved in this transformation were identified and their roles were established by *in vitro* characterization. Detailed mechanistic studies revealed a novel flavin dependent enzyme that catalyzes superoxide mediated H-atom abstraction from substrate riboflavin (**1**). Structural studies demonstrate that this novel reactivity of RcaE is a consequence of the orientations of substrate RBF (**1**) and cofactor FMN (**2**) in the active site such that formation of C_{4a}-flavin(hydro)peroxide is sterically hindered.

Recently, the whole genome sequence of *Devosia riboflavina* was deposited in the NCBI database.[63] BLAST analysis of RcaE from *Microbacterium maritypicum* with the whole genome sequence revealed the ortholog from *Devosia riboflavina* (Figure 3.9). The neighboring genes of RcaE that are involved in riboflavin catabolism where also found in *D. riboflavina*. This suggests that the riboflavin catabolic pathway identified in this study is identical to the one present in *D. riboflavina*. Thus, ~72 years after the first report of riboflavin catabolism, the gene cluster involved in this transformation has been identified and characterized. This study paves the way for further studies on alternate riboflavin catabolic pathways, which were previously reported but the strains were lost.

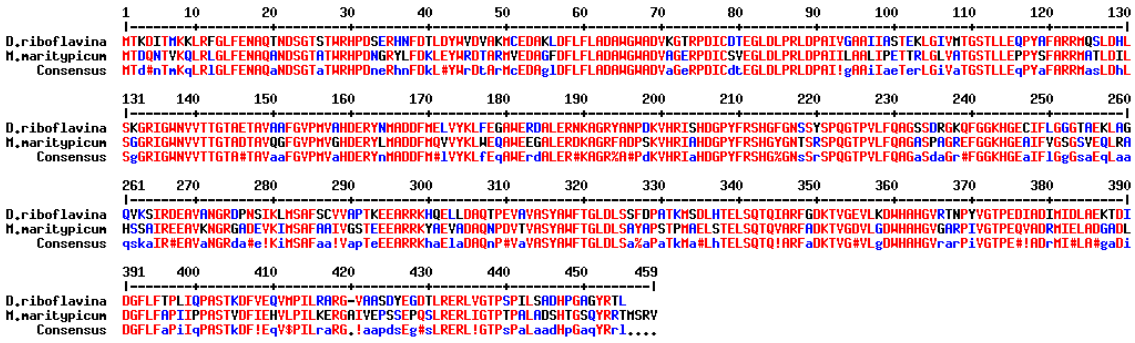


Figure 3.9: Sequence alignment of riboflavin lyase enzyme from *Microbacterium maritypicum* and *Devosia riboflavina*.

3.5 Experimental procedures

Materials: All chemicals were purchased from Sigma-Aldrich (now MilliporeSigma) unless specified otherwise. Nutrient Broth (NB) was from Difco. M9 media was supplied by VWR. Kanamycin was from Teknova and IPTG was obtained from Lab scientific Inc. HPLC and LC-MS solvents were purchased from EMD and were used without further purification. Histrap column (5 ml) was obtained from GE healthcare. Econo-Pack 10 DG desalting columns were purchased from Bio-Rad Laboratories. Large cultures were grown and overexpressed in 2.5 L baffled ultra yield flasks from Thomson Instrument Company.

Growth studies on Microbacterium maritypicum G10: The strain was grown overnight in NB media at 30 °C. 100 µl of this seed culture was added to a 50 ml M9 media (with or without NH_4Cl) supplemented either with glucose or riboflavin (**1**) in a 250 ml Erlenmeyer flask. The cultures were incubated at 30 °C for ~48 hr with shaking. The amount of riboflavin supplemented was in the range of 1-5 mg/50 ml M9 media.

Growth curves for *M. maritypicum* were obtained by measuring absorbance at 600 nm in UV-Vis spectrophotometer. For HPLC analysis, the spent media was passed through a 10kda filter. For quantitation of lumichrome (**11**) accumulated in the media, 500 µl of media was mixed with 500 µl of 1N KOH solution. The samples were then transferred to quartz cuvettes and UV-Vis spectrum was recorded between 200 – 800 nm. A calibration curve was prepared in a similar manner with authentic standard of lumichrome (**11**). These experiments demonstrated that the lumichrome (**11**) is not degraded further by *M. maritypicum*.

HPLC parameters: An Agilent 1260 HPLC equipped with a quaternary pump was used. The system included a diode array UV-Vis detector and chromatograms were detected using absorbance at 220 nm, 254 nm, 260 nm, 280 nm, 340 nm, 370 nm and 450 nm. Analysis was performed ZORBAX Eclipse XDB-C18 column (15 cm x 4.6 mm, 5 µm particles, Agilent Technologies). Data was processed using ChemStation ver. B.04.01 SP1 (Agilent technologies).

HPLC conditions*:*

A- Water

B- 100 mM Potassium phosphate buffer, pH 6.6

C- Methanol

HPLC method*:* (Flow rate = 1 ml/min)

0 min – 10% A 90% B, 2 min – 10% A 90% B, 12 min – 15% A 10% B 75% C, 17 min – 15% A 10% B 75% C, 18.5 min - 90% A 10% B, 20 min - 90% A 10% B, 25 min - 90% A 10% B.

Orthologs for crystallization of RcaE: Initial crystallization trials did not yield any hits for crystals of RcaE with the enzyme from *M. maritypicum*. Thus, following orthologs of RcaE were identified using BLAST search and gene synthesized from GenScript. All of these genes are currently annotated as Hypothetical proteins. *Herbiconioux sp*. YR403 (71% identity), *Marine actinobacterium* PHSC20C1 (70% identity), *Devosia riboflavina* (67% identity), *Leifsonia aquatic* (72% identity) and *Kitasatospora setae* (68% identity). These genes were cloned in pTHT vector (modified pET28b vector with TEV protease cleavage site after His-tag at the N-terminus).

Overexpression and purification of enzymes: The respective plasmids were transformed into BL21(DE3) competent cells of *E.coli* using electroporation technique. Starter cultures were grown overnight in LB media containing kanamycin (40 μg/ml). 30 ml of this culture was added to 3 L LB media (2 x 1.5 L flasks) with antibiotics and grown at 37 °C with shaking (180 rpm) till OD_{600}~0.6. The flasks were then incubated at 4 °C for ~2 hr without shaking. Cultures were then induced with 500 μM IPTG followed by incubation at 15 °C for ~14 hr with shaking at 180 rpm. The cells were harvested by centrifugation at 9,000 rpm for 10 min and stored in liquid nitrogen until further use. Typical yield was ~9 gm of cell pellet (wet weight) from 3 L cell culture.

For purification, the cell pellets were thawed and resuspended in 40-50 ml of lysis buffer (100 mM Tris.HCl, 150 mM NaCl, pH 7.5) at room temperature in the presence of lysozyme (6-8 mg). The suspension was stirred for ~1.5 hr on an ice-bath and further sonicated to lyse the cells. Cell debris was removed by centrifugation at 18,000 rpm for 20min and the lysate was filtered using 0.22 μm filters. The filtered lysate was loaded onto a His-trap column pre-equilibrated with lysis buffer. The column was washed with 10 column volumes of wash buffer (100 mM Tris.HCl, 20 mM imidazole, 150 mM NaCl, pH 7.5). The protein was then eluted from the his-trap column with elution buffer (100 mM Tris.HCl, 250 mM imidazole, 150 mM NaCl, pH 7.5). Protein containing fractions were identified using Bradford reagent. They were pooled together and concentrated using 15 ml 10kda filters. The concentrated protein was buffer exchanged to 50 mM Tris.HCl, 10% glycerol, pH 7.5 using the Econo-Pac 10DG desalting column.

After desalting RcaE, purified TEV protease was added to it in the ratio TEV:CmoJ = 1:10 and incubated at 30 °C for 4 hr. The solution of proteins was passed through Ni-NTA column, which was pre equilibrated with (100 mM Tris, 150 mM NaCl, pH 7.5). Histag cleaved RcaE was eluted from the column using 100 mM Tris, 20 mM imidazole, 150 mM NaCl, pH 7.5 and further desalted into 10 mM Tris, pH 7.5 buffer. The enzyme was used directly for setting up crystallization trials. The enzyme at -80 °C yielded poor crystals. Protein concentration was determined by measuring the absorbance at 280 nm (A_{280}) and utilizing the extinction coefficient calculated using ProtParam tool of the ExPASy proteomics server. The function of all these orthologs

was confirmed by carrying out biochemical assays. All purified enzymes used riboflavin (**1**) as a substrate in presence of RcaB (reductase), FMN and NADH. Ribose (**12**) and lumichrome (**11**) were formed as products.

Crystallization of RcaE orthologs: Crystallization trials were setup using commercial kits and hanging drop diffusion method. Enzyme concentration was kept in the range of ~13-15 mg/ml. Few hits were obtained with the *Herbiconioux sp*. YR403 ortholog and were optimized further optimized to yield better quality crystals.

Apo RcaE: 2.5 µl of reservoir solution was added to 2.5 µl of purified RcaE on the coverslip and sealed on top of a reservoir containing 100 mM BisTris-HCl pH 6.0 and 1.7 M $(NH_4)_2SO_4$ at 25°C. Rod shaped crystals were obtained after one week which grew larger for ~15 days.

RcaE with FMN: A stock solution of 15 mg/ml protein with 2 mM FMN in 10 mM Tris-HCl, pH 7.5 was prepared. Crystallization was setup at 4 °C in a 24-well plate. The reservoir solution contained 100 mM BisTris-HCl pH 6.5, 45% Polypropylene glycol P400. Diamond shaped crystals appeared after a week which grew larger for ~15 days.

RcaE with RBF and FMN: The solubility of riboflavin in aqueous solution is quite low (~200uM). Co-crystallization and soaking experiments with the above-mentioned conditions did not yield a riboflavin bound structure. To obtain a concentrated solution of riboflavin, 3 mM FMN was dissolved in 10 mM Tris-HCl pH

7.5 buffer and treated with Calf-intestinal Phosphatase (CIP) at 37 °C. After ~2 hr, CIP was inactivated by heating at 65 °C for 20 min. This strategy resulted in 3 mM stock solution of riboflavin. This was used to prepare the final stock solution of 15 mg/ml protein, 2 mM FMN and 2 mM RBF in 10 mM Tris-HCl, pH 7.5. Crystallization was setup with reservoir containing 100 mM BisTri-HCl pH 6.0 and 1.9 M $(NH_4)_2SO_4$ at 25°C. Thick rod shaped crystals were obtained after four days which grew larger for ~15 days.

Glycerol was used as a cryoprotectant while preparing samples for X-ray diffraction.

Collection of X-ray diffraction data: All X-ray diffraction data was collected at the Advanced Photon Source (APS), Argonne National laboratory. Crystals were diffracted at 100 K, at the 23-ID-D beamline using the PILATUS3 6M detector and a wavelength of λ= 1.033 Å. Diffraction data was collected for 240-360 frames and was recorded in a shutterless mode at the rate of one degree/sec. Datasets were integrated and scaled using iMosflm.

Structure determination and refinement: The structure of RcaE bound with FMN was determined by molecular replacement using published structure of LadA PDB id 3B9O (40% sequence similarity with RcaE) as the starting model. The structures were refined iteratively using programs like PHENIX and COOT. The model thus generated was used as a starting model for RcaE (apo form) and RcaE bound with RBF and FMN. The final refinement statistics are summarized in Table 3.2.

4. MECHANISTIC STUDIES ON THE REACTION CATALYZED BY TRYPTOPHAN LYASE (NOSL)*

4.1 Introduction to radical S-adenosyl-L-methionine (SAM) enzymes[64]

Biological transformations often involve radical mechanisms.[65] Some of the most intensively studied systems include metalloenzymes like cytochrome P450's[66] and alkane mono-oxygenases,[67] ribonucleotide reductase,[68] flavoenzymes[69] and vitamin B_{12} dependent enzymes.[70] In 2001, enzymes utilizing [4Fe-4S] cluster and S-adenosylmethionine (SAM) that catalyze diverse radical reactions were classified as a superfamily.[10] With rapid progress in sequencing of bacterial genomes,[71] today over 165,000 enzymes are classified into this superfamily. Almost all known radical SAM enzymes (except Dph2[72] and glycerol dehydratase activating enzyme[73]), utilize S-adenosyl-L-methionine (**4**, SAM) as a source to generate the 5′-deoxyadenosyl radical (**6**, Ado•). The 5′-deoxyadenosyl radical (**6**) abstracts a H-atom either from substrate or enzyme to initiate the desired transformation. Radical **6** is also found in various vitamin B_{12} dependent enzymes and is present across all kingdoms of life.[74]

---

*Reprinted in parts with permission from "Tryptophan Lyase (NosL): Mechanistic Insights from Substrate Analogues and Mutagenesis" by Bhandari, D. M., Xu, H., Nicolet, Y., Fontecilla-Camps, J. C., & Begley, T. P. *Biochemistry*, **2015**, *54*(31), 4767-4769. Copyright 2015 American Chemical Society and from "Mechanistic Studies on Tryptophan Lyase (NosL): Identification of Cyanide as a Reaction Product" by Bhandari, D. M., Fedoseyenko, D., & Begley, T. P. *J. Am. Chem. Soc.* **2017**, (manuscript submitted). Copyright 2017 American Chemical Society.

Most enzymes in radical SAM superfamily have a characteristic CX_3CX_2C motif (Figure 4.1B).[12] The three-cysteine residues coordinate with three of the four irons in a [4Fe-4S] cluster within the active site of these enzymes. SAM binds through the amine and carboxylate of the methionine moiety to the fourth iron of the [4Fe-4S] cluster. The cluster is sensitive to oxygen and in absence of a reducing agent; the cluster is in $[4Fe\text{-}4S]^{+2}$ form. In presence of a suitable reducing agent, the cluster can be reduced to its catalytically active $[4Fe\text{-}4S]^{+1}$ form. As shown in Figure 4.1A, the reduced cluster transfers one electron to the sulfonium center of SAM, which leads to homolytic cleavage of the S−5′C bond generating 5′-deoxyadenosyl radical (**6**, Ado•). In most cases, **6** initiates chemistry on the substrate and is converted to a product in a controlled fashion. The diverse chemistry catalysed by this family of enzymes signifies the strategy evolved by nature to fine tune the reactivity of unstable radical species and obtain products that would be otherwise inaccessible.

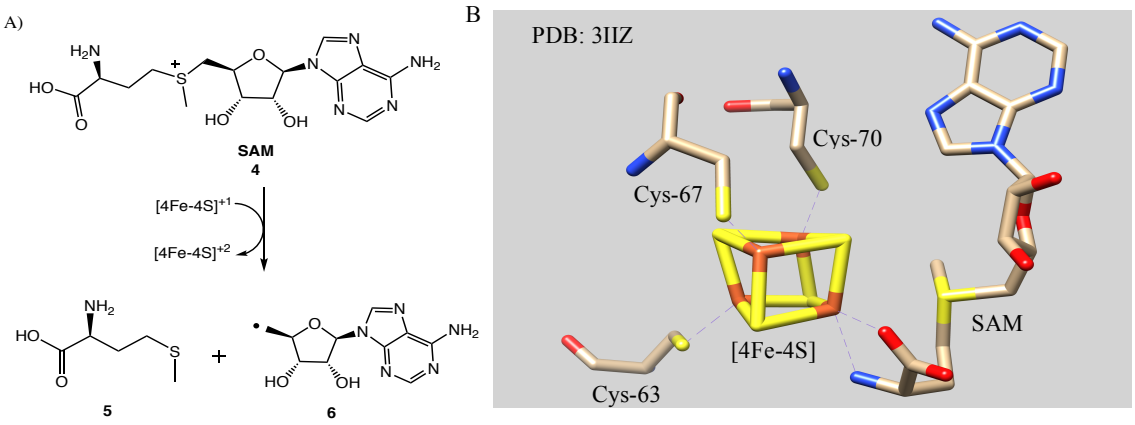


Figure 4.1: Introduction to the radical SAM enzymes. A) Reductive homolytic cleavage of SAM (**4**) leads to formation of the 5′-deoxyadenosyl radical (**6**) and methionine (**5**). B) Typical active site of a radical SAM enzyme showing binding of cysteines and SAM to the irons of the [4Fe-4S] cluster.

4.2 Aromatic amino acid lyases subfamily[13]

A subgroup (>1000 annotations by bioinformatics) in radical SAM superfamily consists of enzymes that generate fragments of aromatic amino acids and utilize them in their respective biosynthetic pathways. These include ThiH from prokaryotic thiamin biosynthetic pathway, HydG involved in maturation of H-cluster of hydrogenase, F_0 synthase or CofG/CofH from F_{420} biosynthesis and NosL from nosiheptide biosynthesis. As shown in Figure 4.2, ThiH, HydG and CofH utilize L-tyrosine (**78**) as a substrate while NosL uses L-tryptophan (**13**) as the substrate. ThiH catalyzes the formation of dehydroglycine (**80**), which gets incorporated into thiamin (**81**). HydG generates the cyanide (**82**) and carbon monoxide (**83**) ligands of the [Fe-Fe]-hydrogenase H-cluster (**84**). The p-cresol radical (**85**) generated by CofH is incorporated into cofactor F_{420} (**86**). NosL is involved in biosynthesis of 3-methyl-2-indolic acid (**15**), which is further incorporated into a thiopeptide antibiotic nosiheptide (**89**).

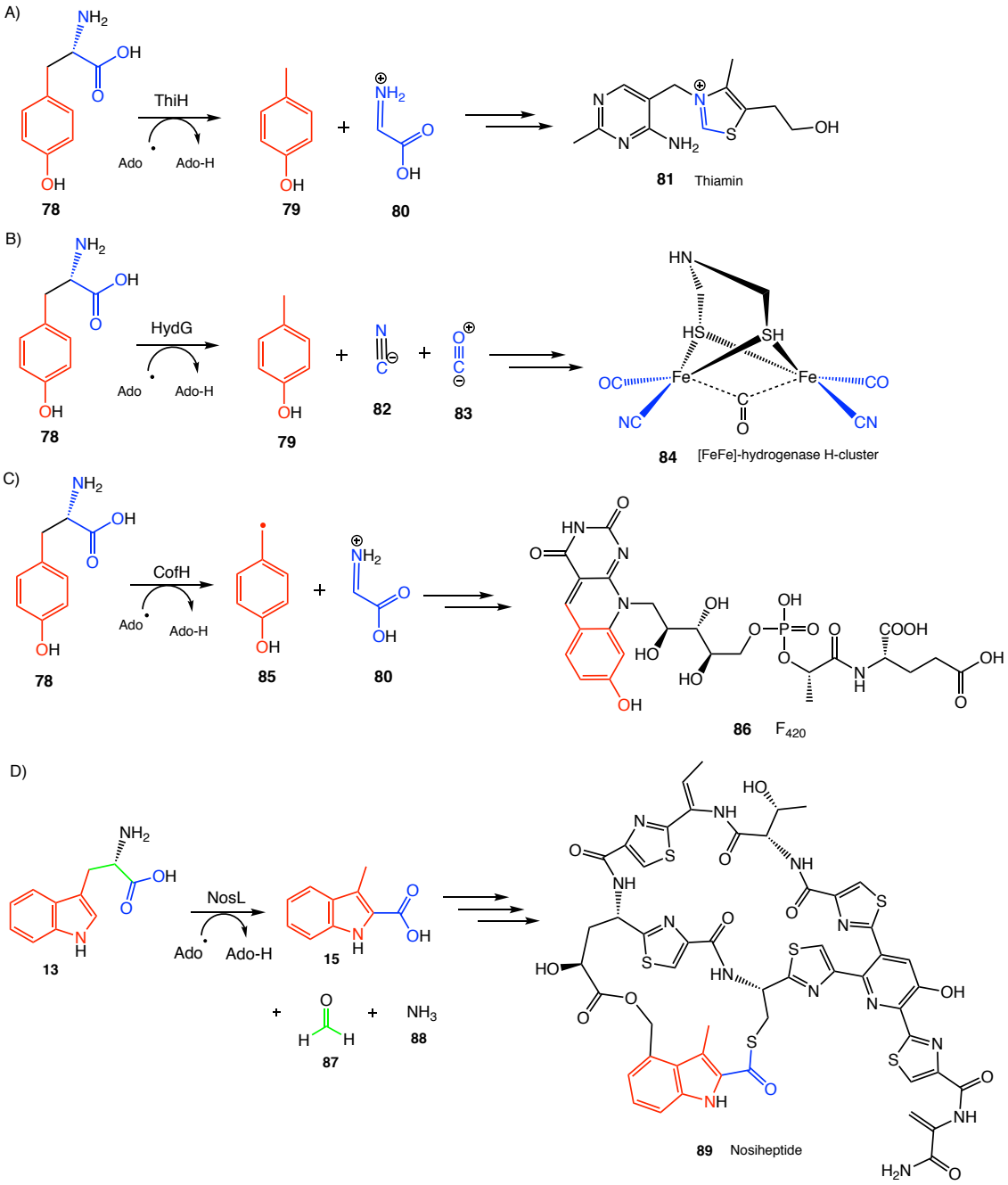


Figure 4.2: Examples of the aromatic amino acid lyases from radical SAM superfamily. A) ThiH catalyzes formation of dehydroglycine (**80**) which is incorporated into thiamin (**81**). B) HydG catalyzes formation of carbon monoxide (**83**) and cyanide (**82**) fragments which are incorporated into [Fe-Fe]-hydrogenase H-cluster (**84**). C) In CofH, the p-cresol radical (**85**) is incorporated into cofactor F_{420} (**86**). D) NosL catalyzes formation of 3-methyl-2-indolic acid (**15**), a fragment of nosiheptide (**89**).

4.3 Background

Nosiheptide (**89**)[75] is a highly modified polythiazolyl heterocyclic peptide antibiotic active against highly resistant pathogenic strains of *Staphylococcus aureus, Streptococcus pneumonia, Clostridium difficile* and several enterococci.[76] The precursor peptide is ribosomally encoded and the final molecule contains five thiazoles, a tetrasubstituted pyridine and an unusual indolic acid (see Figure 4.2D).[77] Recently, the gene cluster involved in the biosynthesis of nosiheptide was identified and NosL was demonstrated to be the enzyme required for catalysis of the remarkable transformation of L-tryptophan (**13**) to 3-methyl-2-indolic acid (**15**).[21]

Early reports on mechanism for this enzyme family proposed that the chemistry is initiated on substrate by abstraction of H-atom (by Ado•) from phenolic O−H for ThiH, HydG and F_0 synthase, and from indolic N−H for NosL (Figure 4.3A).[78] This hypothesis was based on enzymatic reactions done in presence of deuterated substrates (C−H bonds labeled to C−D bonds). The 5′-deoxyadenosine (**93**) thus generated did not show any deuterium label incorporation. Further, when the enzymatic reaction was performed in D_2O buffer, incorporation of deuterium in 5′-deoxyadenosine was observed.[79] However, this only implies that Ado• abstracts H-atom from a heteroatom on substrate in all enzymes of this family.

Recently, crystal structure of NosL was obtained complexed with [4Fe-4S] cluster + SAM + L-tryptophan.[22] As shown in Figure 4.3B, in the active site of NosL, 5′-C of adenosine from S-adenosyl-L-homocysteine (SAH) is 3.7 Å away from N-atom of amino group of L-tryptophan (**13**). This indicates that H-atom is abstracted from

nitrogen-atom of amino group instead of nitrogen-atom from indole ring of **13**. Crystal structure of NosL is the first direct evidence of H-atom abstraction from the amino group of a substrate by the 5′-deoxyadenosyl radical (**6**) in the radical SAM superfamily.

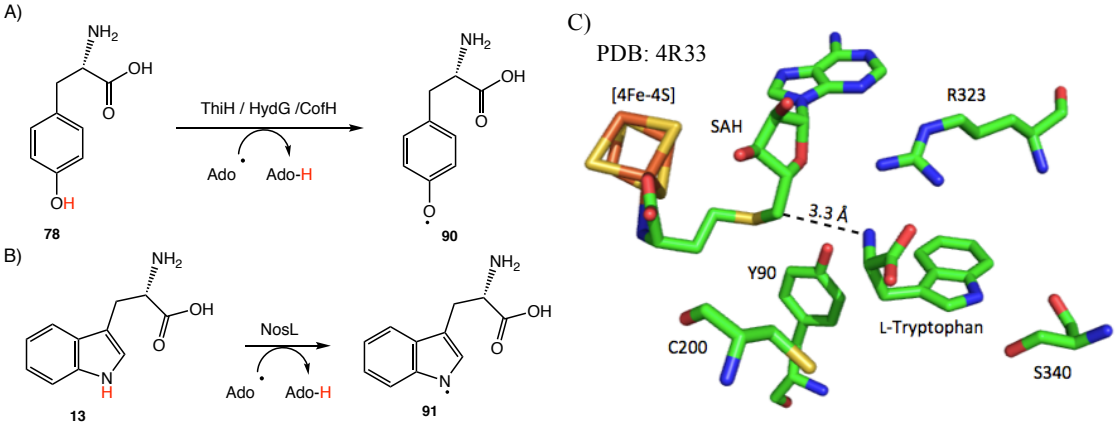


Figure 4.3: Early mechanistic studies on aromatic amino acid lyases. A) For tyrosine lyases it was proposed that the 5′-deoxyadenosyl radical (**6**) abstracts H-atom from the phenolic O-H group generating **90**. B) In NosL catalyzed reaction, the 5′-deoxyadenosyl radical (**6**) was proposed to abstract a H-atom from indolic N-H of L-tryptophan generating **91**. C) Active site architecture of NosL shows that the C_5' of SAM is close to the amino N-H instead of indolic N-H of L-tryptophan (**13**).

4.4 Results and Discussion

4.4.1 *In vitro* reconstitution of Tryptophan Lyase (NosL)

When the NosL catalyzed reaction was carried out using L-tryptophan (**13**) as substrate and dithionite as the reductant, HPLC analysis showed primarily the formation of 3-methyl indole (**92**) while 3-methyl-2-indolic acid (**15**) was produced in small quantities and was characterized using LC-MS (Figure 4.4). Using excess dithionite

reduced the amount of 3-methyl-2-indolic acid (**15**) generated in comparison to 3-methyl indole (**92**). The higher ratio of **15** to **92**, was obtained by lowering the amount of dithionite to 1-2 equivalents of the enzyme concentration.

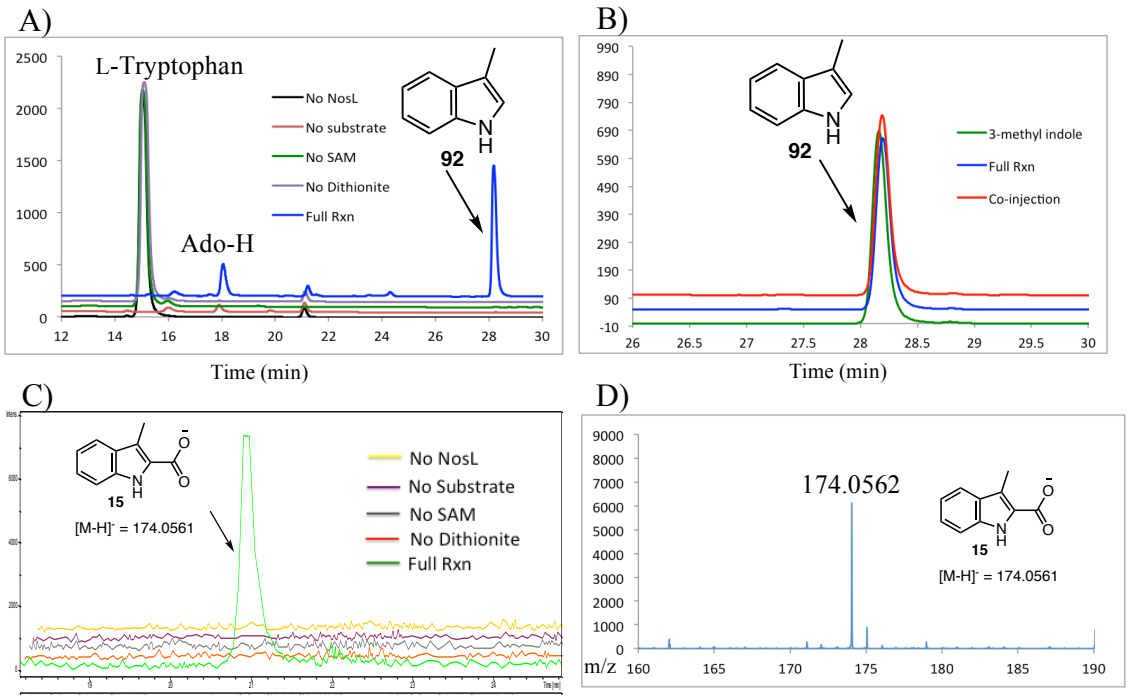


Figure 4.4: Analysis of the NosL-catalyzed reaction with L-tryptophan. A) HPLC chromatogram at 288 nm showing consumption of L-tryptophan (**13**) and formation of 5′-deoxyadenosine (Ado-H) and 3-methylindole (**92**) when all components of the enzymatic reaction were present. B) HPLC chromatogram of authentic standard of 3-methylindole (**92**) and its co-injection with the 'full reaction' sample. C) LC-MS analysis showing the extracted ion chromatogram (EIC) $[M-H]^-$ = 174.0 Da indicating formation of 3-methyl-2-indolic acid (**15**) only when all components of the enzymatic reaction were present. D) MS of 3-methyl-2-indolic acid (**15**) generated in the 'full reaction' sample.

However, when the NosL-catalyzed reaction with L-tryptophan (**13**) was performed in the presence of the flavodoxin system (FldA/FldR), HPLC analysis revealed that in the 'full reaction' sample, 3-methyl-2-indolic acid (**15**) was generated as the major product while 3-methylindole (**92**) was produced in small quantities. As shown in Figure 4.5, formation of 3-methyl-2-indolic acid (**15**) was further confirmed by LC-MS and co-injection with its authentic standard.

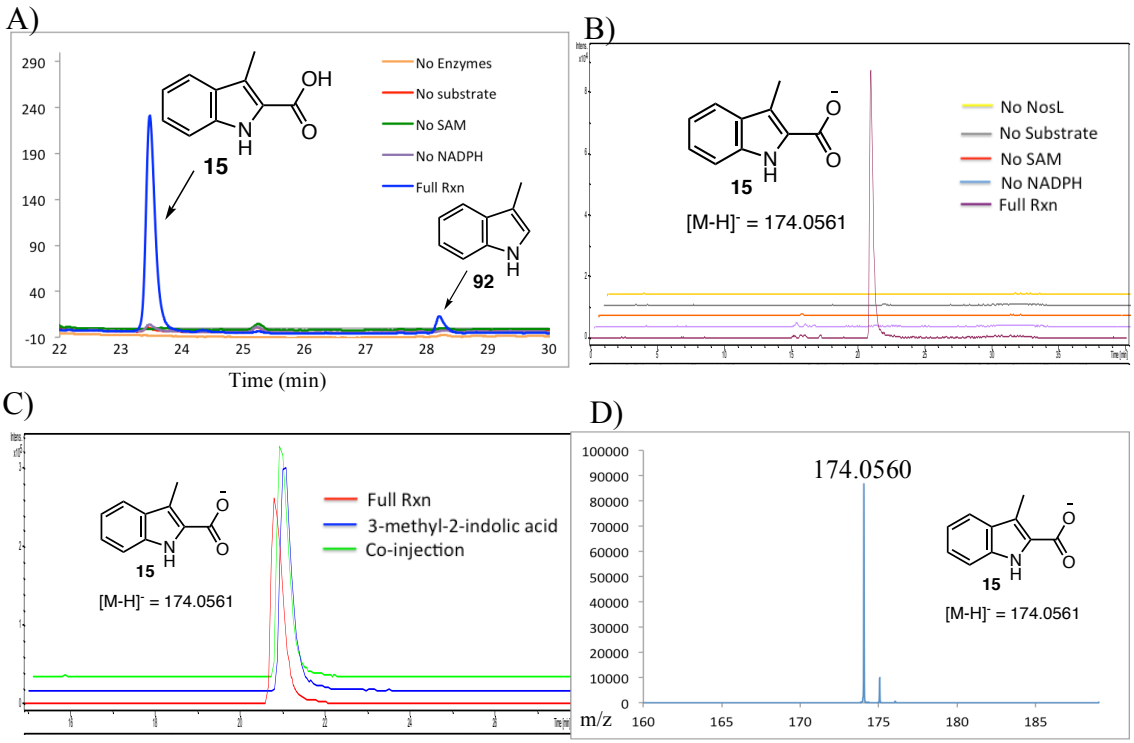


Figure 4.5: Analysis of the NosL-catalyzed reaction with L-tryptophan (**13**) in the presence the FldA/R. A) HPLC analysis at 288 nm showing the formation of 3-methyl-2-indolic acid (**15**) and 3-methylindole (**92**) in the 'full reaction' sample. B) LC-MS analysis showing EIC $[M-H]^-$ = 174.0 Da demonstrating the formation of 3-methyl-2-indolic acid (**15**) only when all components of the enzymatic reaction are present. C) LC-MS analysis showing EIC $[M-H]^-$ = 174.0 Da confirming the formation of **15** by co-injection with the authentic standard. D) MS of 3-methyl-2-indolic acid (**15**) generated in the 'full reaction' sample.

4.4.2 Position of the H-atom abstraction in the NosL catalyzed reaction[80]

The structure of NosL suggests that the 5′-deoxyadenosyl radical (**6**) abstracts a H-atom from the amino rather than from the indolic NH of tryptophan as previously proposed. To test this prediction in a biochemical assay, the NosL reaction was run using a tryptophan analog **94** in which the indole NH is replaced by sulfur. This reaction generated **95** and **96** as products consistent with hydrogen atom abstraction from the amino group (Figure 4.7). Another study using the benzofuran analog of tryptophan gave similar results.[81]

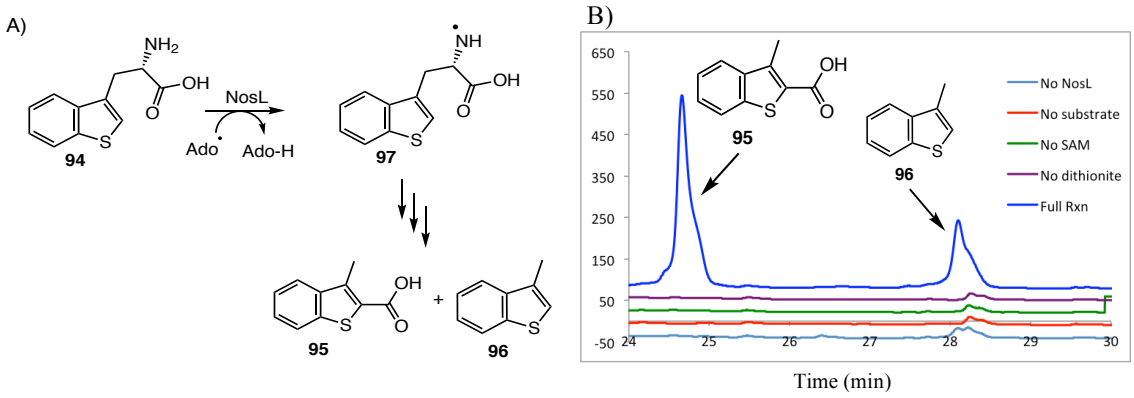


Figure 4.7: Analysis of the NosL enzymatic reaction with L-3-benzothienylalanine (**94**). A) Scheme showing products of reaction with **94** as substrate. B) HPLC analysis of reaction mixture at 254 nm revealed formation of compounds **95** and **96** only in the presence of all the components. C) HPLC analysis showing co-injection of the authentic standard of 3-methylbenzothiophene (**96**) with the compound generated in the 'full reaction' sample. C) HPLC analysis showing co-injection of the authetic standard of 3-methylbenzothiophene-2-carboxylic acid (**95**) with the compound generated in the enzymatic reaction. C) LC-MS analysis showing EIC $[M-H]^-$ = 191.0 Da confirming formation of **95**. D) MS of **95** generated in the 'full reaction' sample.

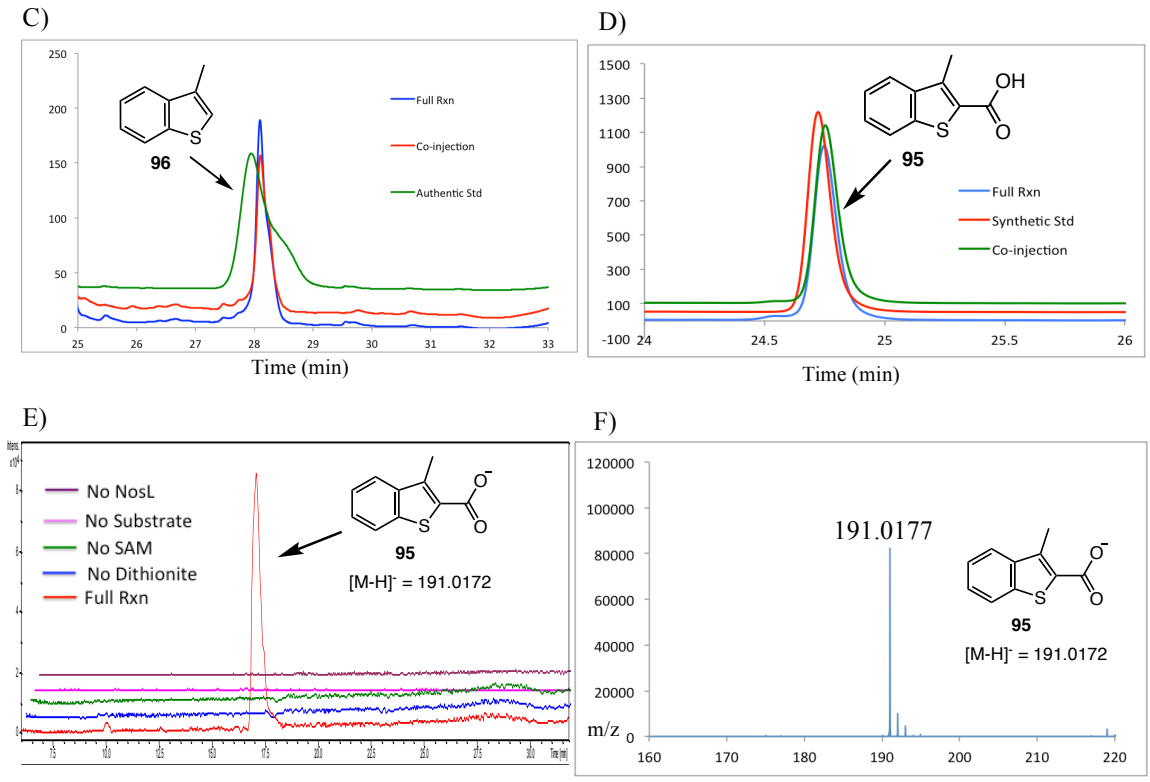


Figure 4.7 continued.

4.4.3 Reaction of NosL in deuterated buffer

To confirm the position of H-atom abstraction from L-tryptophan (**13**) during the NosL catalyzed reaction, the enzymatic reaction was performed in 90% D_2O buffer. As shown in Figure 4.8, LC-MS analysis of the 5′-deoxyadenosine (**93**, Ado-H) showed incorporation of multiple deuteriums from solvent. This observation is similar to reports on other aromatic amino acid lyases previously discussed: HydG, CofH and ThiH (unpublished). This demonstrates that the 5′-deoxyadenosyl radical (**6**) abstracts H-atom from a solvent exchangeable position of the substrate.

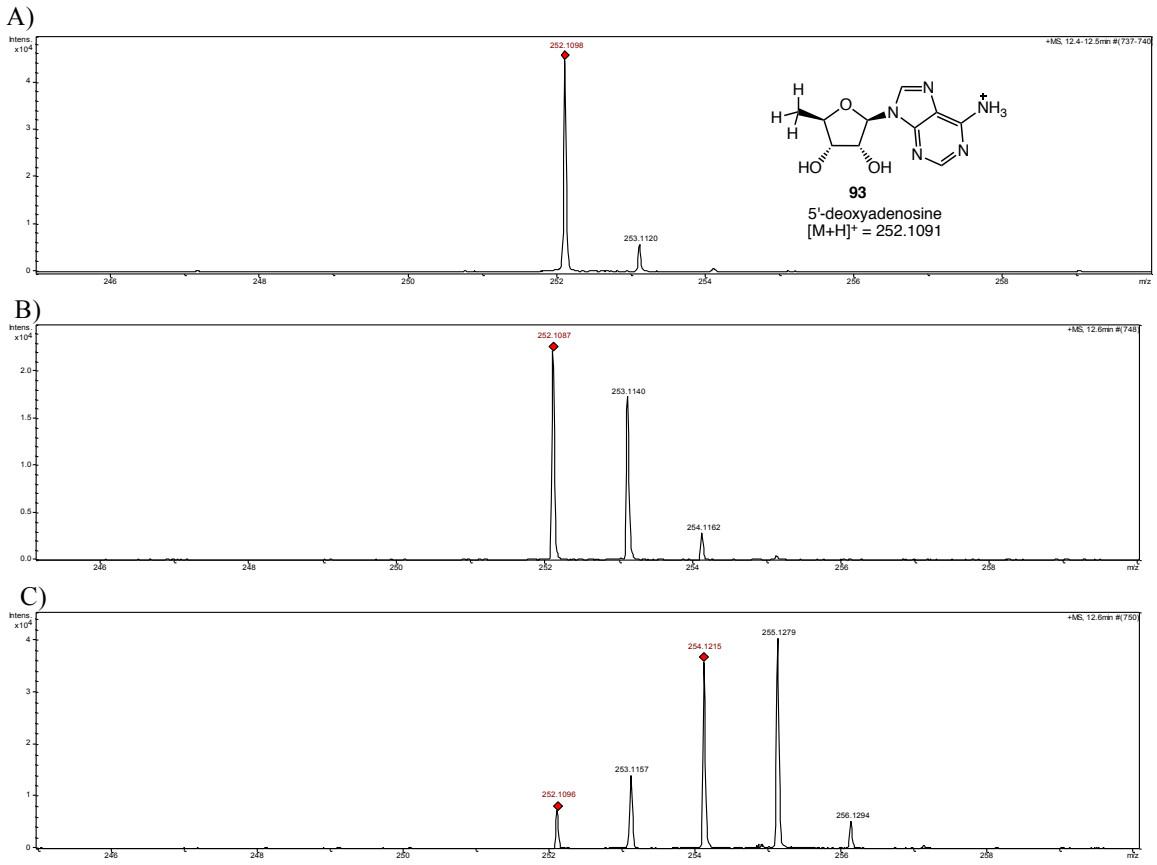


Figure 4.8: MS analysis of deuterium incorporation in 5′-deoxyadenosine (**93**). A) MS of the authentic standard of **93**. B) MS of **93** generated in absence of substrate during the NosL-catalyzed reaction. One deuterium incorporation in **93** was observed. C) MS of **93** generated during NosL catalyzed reaction in presence of L-tryptophan (**13**) as substrate. Incorporation of multiple (upto three) deuteria in **93** was observed.

When the enzymatic reaction is carried out in deuterated buffer, all the solvent exchangeable protons will be replaced by deuterium. As shown in Figure 4.9, the 5′-deoxyadenosyl radical (**6**, Ado•) can abstract a deuterium atom from the amino group of the substrate thereby incorporating a deuterium into 5′-deoxyadenosine (**93**). The nitrogen-centered radical can further abstract a H-atom from 5′-deoxyadenosine. This hydrogen on the amino group can be exchanged with the deuterium present in the buffer.

The 5′-deoxyadenosyl radical (**6,** Ado•) can abstract another deuterium atom from the amino group and so on. Thus the reversibility of H-atom abstraction between the substrate and 5′-deoxyadenosyl radical (Ado•) leads to the incorporation of multiple deuteria into 5′-deoxyadenosine when the enzymatic reaction is carried out in the deuterated buffer.[82]

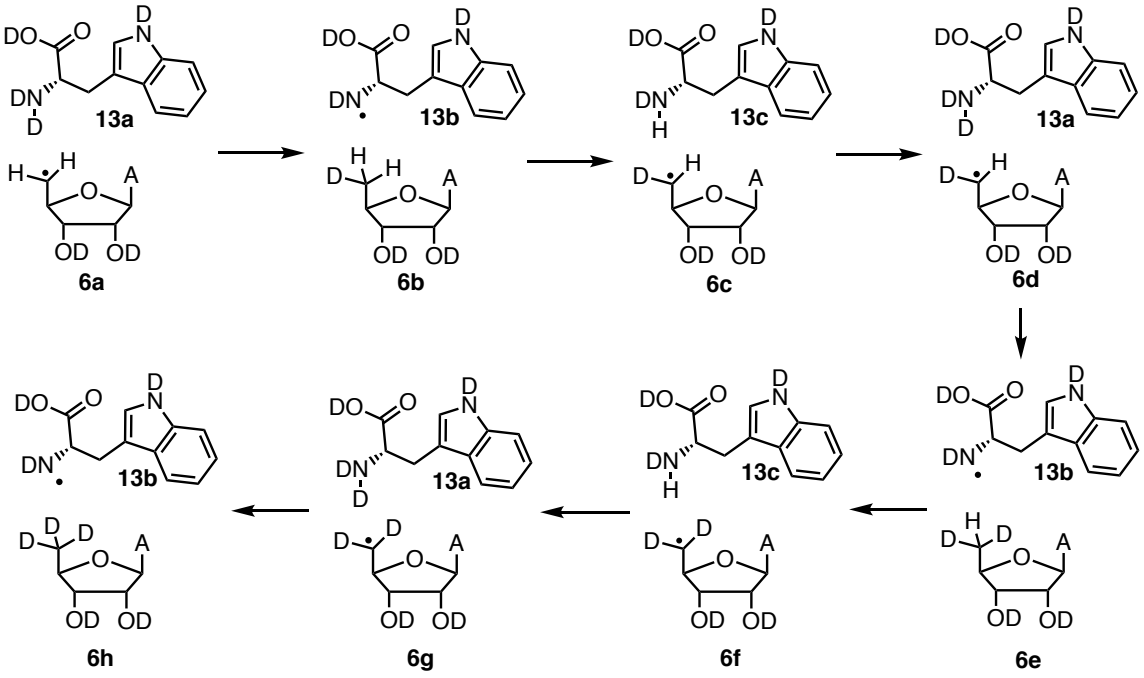


Figure 4.9: Mechanistic proposal for the NosL-catalyzed multiple deuterium incorporation in 5′-deoxyadenosine (Ado-H). [A = Adenine]

4.4.4 Mechanistic proposal for the NosL catalyzed reaction[80]

Based on the previous studies and some preliminary biochemical studies discussed above, a mechanistic proposal for the NosL catalyzed reaction is shown in Figure 4.10. In this mechanism, the 5′-deoxyadenosyl radical (**6**, Ado•) abstracts a H-

atom from the amino group of **13** generating **14**. β-bond scission of the C_α-C_β bond generates radical **98** and dehydroglycine (**80**). Addition of **98** to **80** gives **99**. A second β-bond scission generates **100,** which then isomerizes to the product **15**. The byproduct, radical **102**, may abstract a hydrogen atom from an unidentified source to form imine **103** which then undergoes hydrolysis to formaldehyde (**87**). This mechanism is supported by labeling experiments identifying the origin of key atoms in the product, by the trapping of radical **98** to give **92** and by a recent structure of the enzyme substrate complex, which suggests formation of **14** by H-atom abstraction instead of **91**.

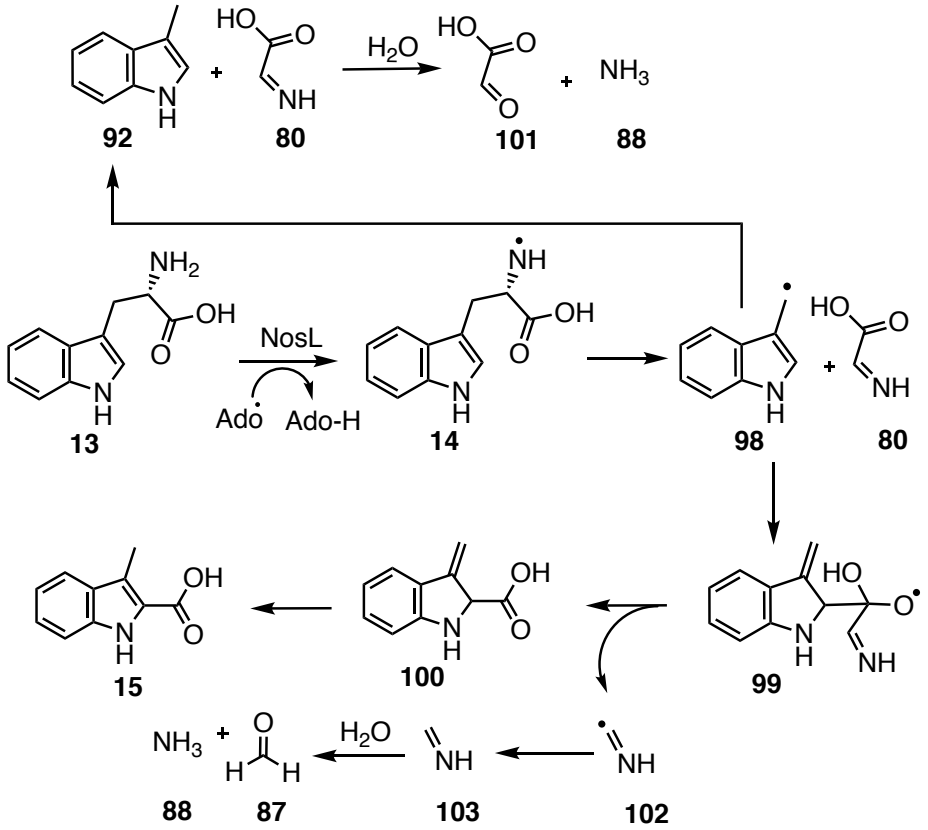


Figure 4.10: Mechanistic proposal of the NosL-catalyzed reaction involving C_α-C_β scission.

4.4.5 Regiochemical control of the NosL catalyzed reaction

The amino centered radical of tryptophan could in principle undergo three possible beta scission reactions involving decarboxylation, deprotonation or C_α-C_β bond cleavage. The crystal structure suggests that hydrogen bonding between the amino group and Arg 323 is likely to control the regiochemistry of this fragmentation reaction (Figure 4.3C). To test this, we examined the reaction of NosL with D-tryptophan (**104**). With this compound, the amino group and the proton at Cα are interchanged thus removing the controlling hydrogen bond to Arg 323. When D-tryptophan (**104**) is used as a substrate, all three possible β-scission reactions occur to give 3-methylindole (**92**, C_α-C_β bond cleavage), indole-3-pyruvic acid (**111**, C_α-H bond cleavage) and indole-3-acetaldehyde (**108**, decarboxylation) (Figures 4.11 – 4.13). Mechanistic proposals for the formation of these three products is shown in Figure 4.11A and 4.11B.[83]

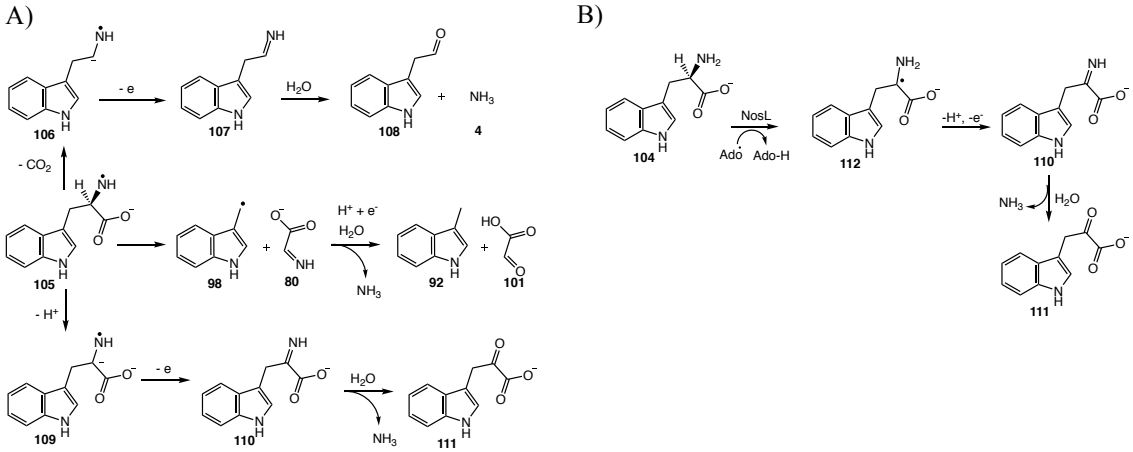


Figure 4.11: Determining regio-control of the NosL catalyzed reaction. A) Three possible β-scission reactions occurring from nitrogen centered radical in D-tryptophan (**104**). B) An alternate pathway for the formation of indole-3-pyruvic acid (**111**).

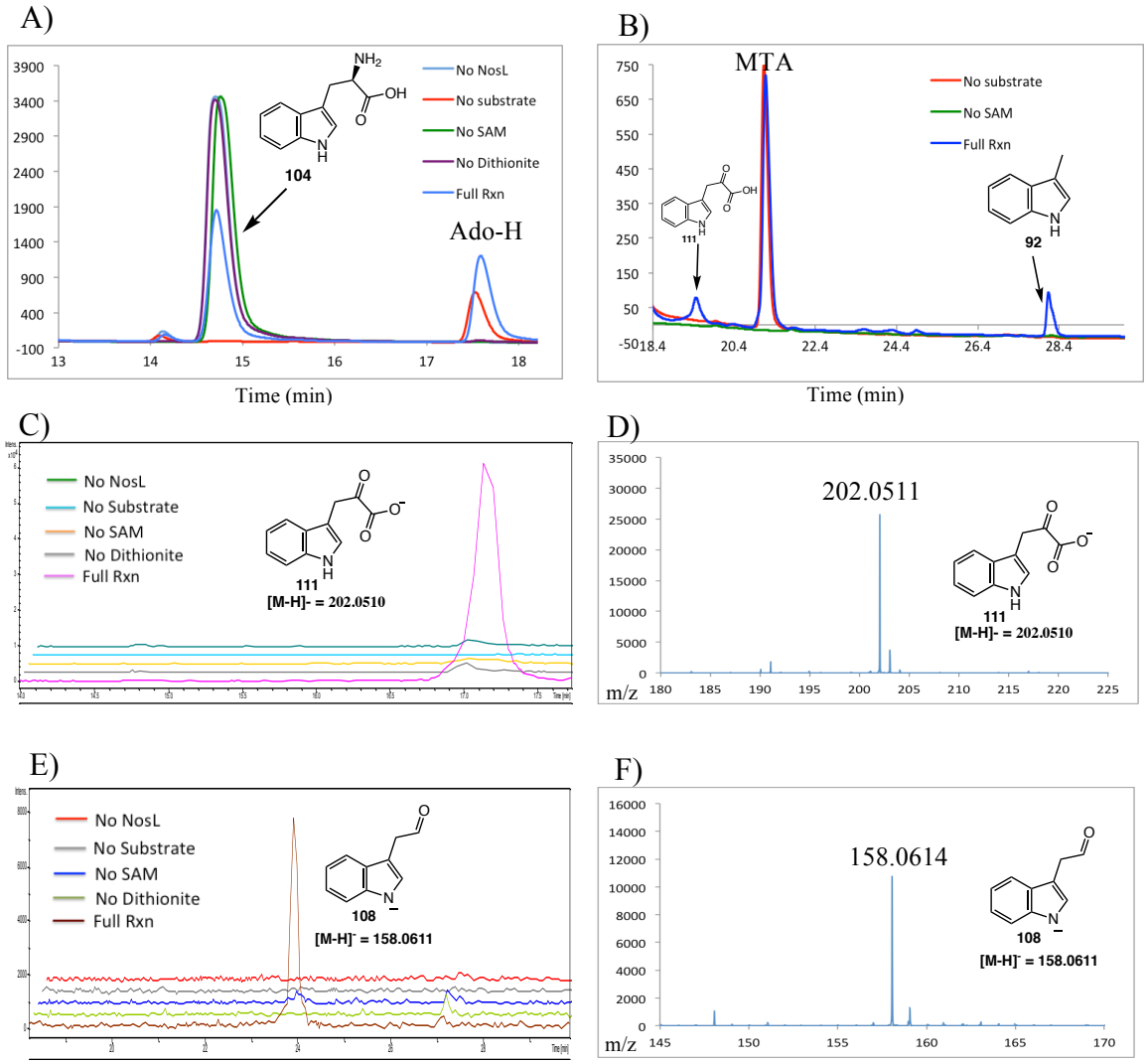


Figure 4.12: Analysis of the NosL reaction with **104**. A) HPLC analysis at 280 nm showing consumption of **104** in the presence of all components of the enzymatic reaction. B) HPLC chromatogram showing formation of 3-methylindole (**92**) and indole-3-pyruvic acid (**111**) in the 'full reaction' sample. MTA is an impurity present in commercial SAM. C) LC-MS analysis showing EIC $[M-H]^-$ = 202.0 Da confirming the formation of **111** in the 'full reaction' sample. D) MS of **111** generated during the NosL-catalyzed reaction. E) LC-MS analysis showing EIC $[M-H]^-$ = 158.0 Da indicating formation of indole-3-acetaldehyde (**108**) in the 'full reaction' sample. F) MS of **108** generated enzymatically.

To confirm the formation of indole-3-acetaldehyde (**108**), 100ul of the 'full reaction' sample was treated with 1mg of $NaBH_4$ and incubated at room temp for ~1 hr. Analysis of this reduced sample by LC-MS showed the disappearance of the signal corresponding to **108** and formation of a new signal corresponding to tryptophol (**113**), Figure 4.13. Formation of **113** was further confirmed by co-injection with its authentic standard.

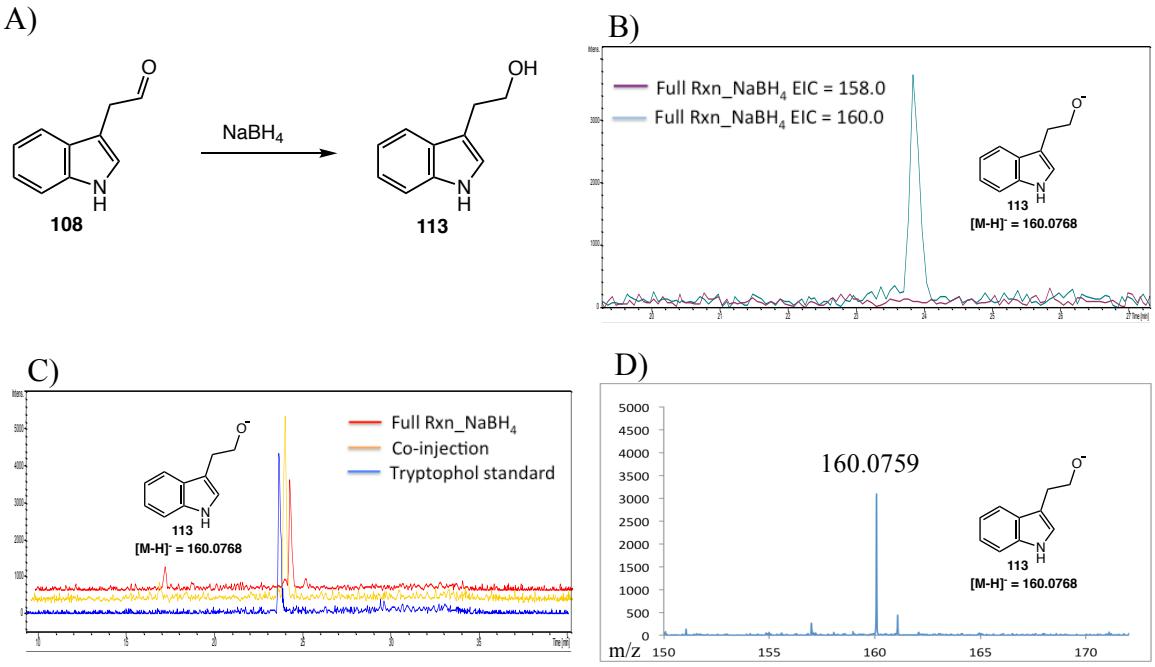


Figure 4.13: Analysis of the tryptophol (**113**) generated by $NaBH_4$ reduction of indole-3-acetaldehyde (**108**) during the NosL reaction with D-tryptophan (**104**). A) Scheme depicting reduction of **108** by $NaBH_4$. B) After treatment of the 'full reaction' sample with $NaBH_4$, LC-MS analysis showed loss of signal EIC $[M-H]^-$ = 158.0 Da and a new signal with $[M-H]^-$ = 160.0 Da was observed. C) LC-MS analysis showing EIC $[M-H]^-$ = 160.0 corresponding to the authentic standard of **113** and its co-injection with $NaBH_4$ reduced 'full reaction' sample. A single peak confirms formation of **113**. D) MS of **113** generated after $NaBH_4$ reduction of the 'full reaction' sample.

4.4.6 Role of active site residues in the NosL catalyzed reaction

The key interactions between tryptophan and the enzyme are shown in Figure 4.3C. The role of all active site residues in catalysis was probed using mutagenesis. Arg 323 is hydrogen bonded to the amino and carboxylate groups of L-tryptophan (**13**). This interaction is likely to play a role in controlling the regiochemistry of the β-scission reaction and in binding the glycine imine for reaction at C_2 of the indole. In support of this, the R323K mutant produces indole-3-pyruvic acid (**111**) instead of **15** and no products arising from either C_α-C_β bond scission or C_α-C(O) bond scission were detected (Figure 4.14).

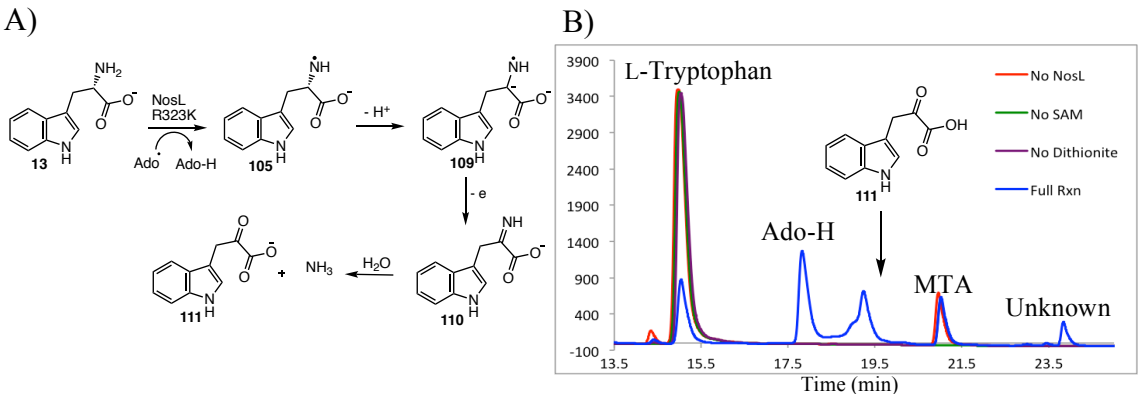


Figure 4.14: Analysis of the enzymatic reaction catalyzed by the NosL R323K mutant with L-tryptophan. A) Mechanistic proposal for the formation of **111** from L-tryptophan. B) HPLC analysis of the enzymatic reaction at 280 nm shows formation of indole-3-pyruvic acid (**111**) and an unidentified new signal only in the 'full reaction' sample. C) LC-MS analysis showing EIC $[M-H]^-$ = 202.0 Da confirming the generation of **111**.

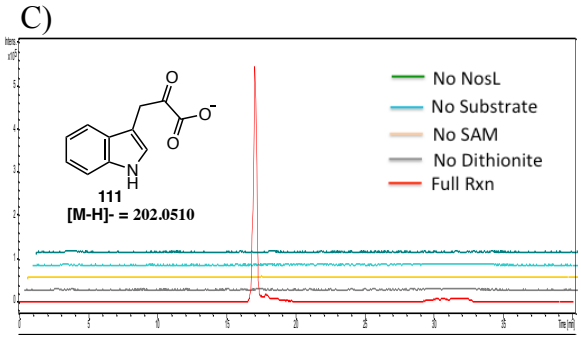


Figure 4.14 continued.

Additional variants of the NosL active site namely Y90F, C200A and S340A were found to generate 3-methyl-2-indolic acid (**15**) from L-tryptophan (**13**) as shown in Figure 4.15. This suggests that these resides are not directly involved in NosL catalysis.

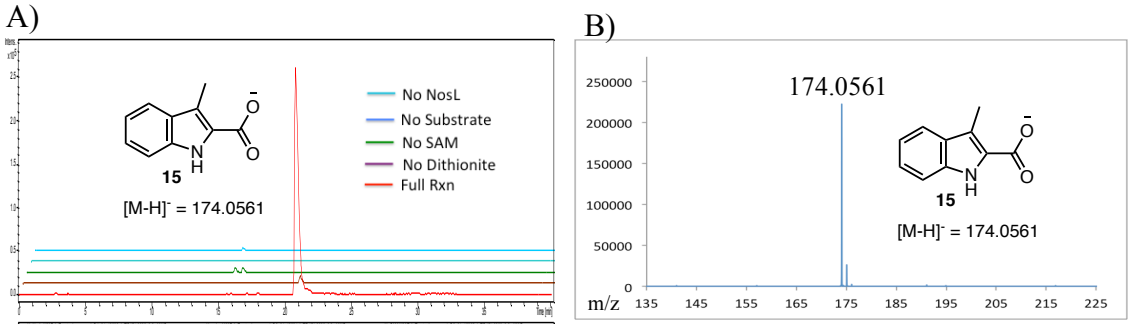


Figure 4.15: Analysis of the NosL active site variants. A) LC-MS analysis of Y90F showing EIC $[M-H]^-$ = 174.0 Da confirming the formation of **15**. B) MS of **15** generated during the enzymatic reaction with Y90F. C) LC-MS analysis of C200A showing EIC $[M-H]^-$ = 174.0 Da confirming the formation of **15**. D) MS of **15** generated during the enzymatic reaction with C200A. E) LC-MS analysis of S340A showing EIC $[M-H]^-$ = 174.0 Da confirming the formation of **15**.

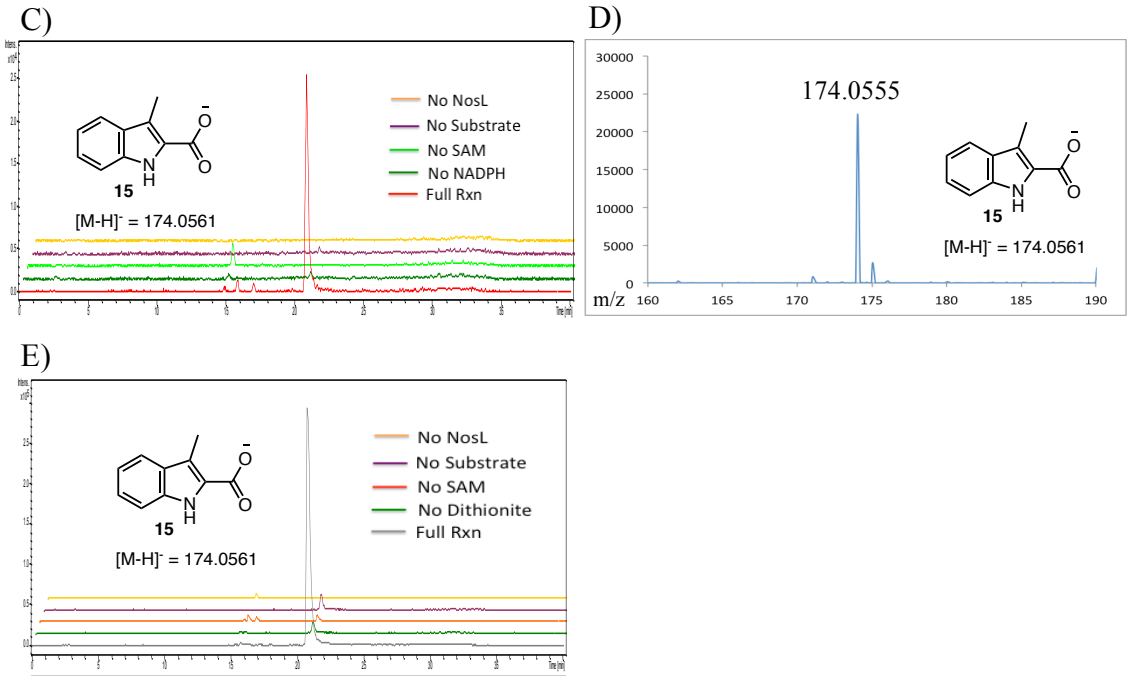


Figure 4.15 continued.

4.4.7 Reaction of NosL with L-tryptophan thiocarboxylate (115)

To gain further insights into the β-scission after H-atom abstraction from the amino group (Figure 4.16), we investigated the reaction of NosL in presence of L-tryptophan thiocarboxylate (**115**). If the reaction proceeds via C_α-C_β bond scission, it will generate 3-methylindole (**92**) while if it proceeds via C_α-C(O) bond scission, it will lead to the formation of indole-3-acetaldehyde (**108**). HPLC analysis of the NosL tryptophan thiocarboxylate reaction mixture showed complete substrate consumption (Figure 4.17) and formation of a new product, which was identified as aldehyde **108** based on LC-MS analysis. This product was also reduced with $NaBH_4$ to form alcohol **113**, which was chromatographically identical with an authentic standard. Remarkably,

3-methylindole (**92**), the major product formed with L-tryptophan, was not detected. Addition of the thioformyl radical (**118**) to 3-methylindole (**92**) is retarded due to stabilization by the sulfur and leads to the trapping of **108**.

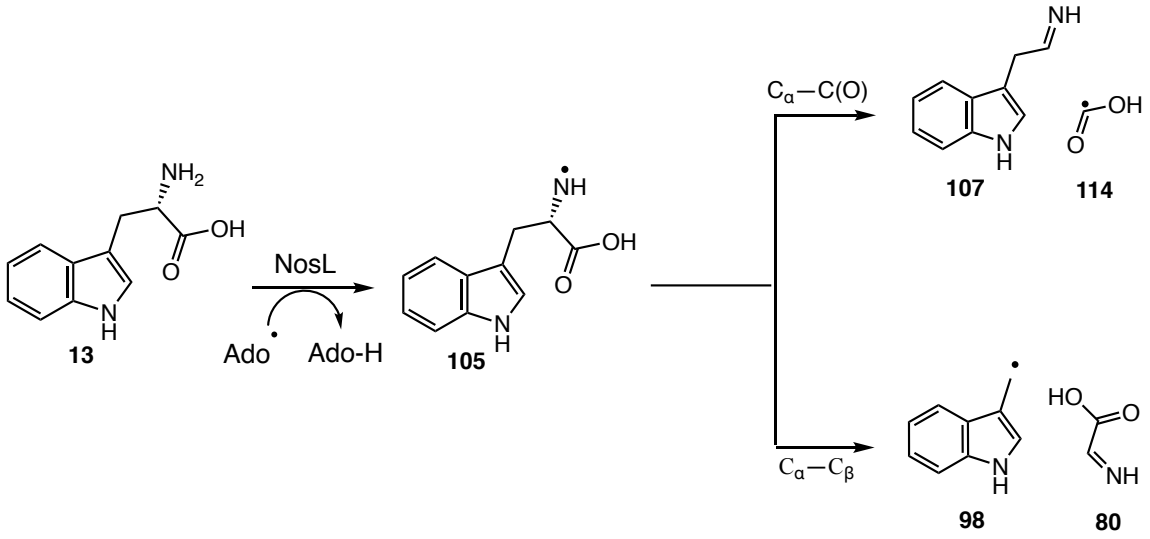


Figure 4.16: After H-atom abstraction from the amino group of L-tryptophan, two possible β-scission reactions will lead to different products.

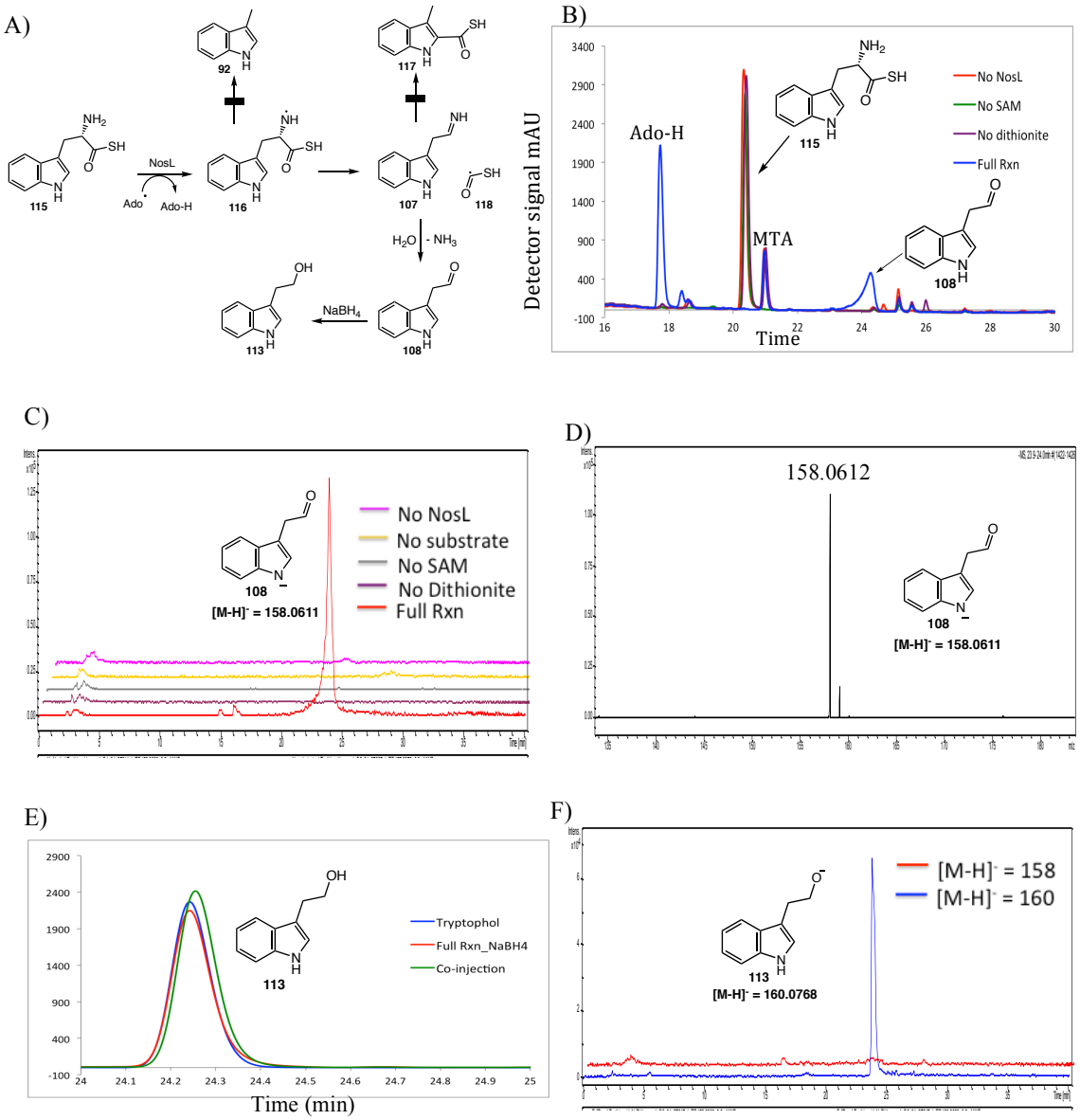


Figure 4.17: Analysis of NosL catalyzed reaction with L-tryptophan thiocarboxylate (**115**). A) Mechanistic proposal for the C_{α}-C(O) bond scission. B) HPLC analysis at 280 nm shows complete consumption of substrate **115** and formation of aldehyde product **108**. C) LC-MS analysis showing EIC $[M-H]^-$ = 158 Da suggesting the formation of **108**. D) MS of reaction product **108** in negative mode. E) Co-injection of **113** with authentic standard confirming its formation after $NaBH_4$ reduction. F) LC-MS analysis of the 'Full Rxn' sample after treating with $NaBH_4$ showing reduction of aldehyde **108** to tryptophol (**113**).

We have previously observed the formation of aldehyde **108** in the reaction of D-tryptophan with NosL but viewed this fragmentation as an off-pathway reaction because of the altered tryptophan stereochemistry. Aldehyde **108** however was not previously detected from L-tryptophan (**13**). Motivated by the results of the tryptophan thiocarboxylate experiment, we repeated the NosL/L-tryptophan reaction (with all variants of NosL) using fluorescence rather than absorbance for product detection to give greater sensitivity for the detection of aldehyde **108**. HPLC analysis of the resulting reaction mixture showed a new signal at 24.3 min (Ex = 270 nm, Em = 352 nm). LC-MS analysis confirmed the identification of this product as the aldehyde **108** (Figure 4.18).

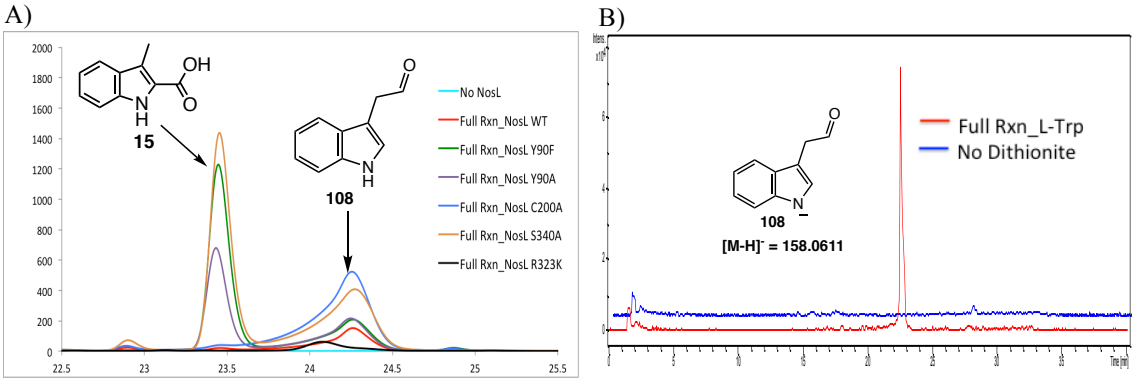


Figure 4.18: Detection of aldehyde **108** in NosL and its variants. A) Fluorescence analysis at Ex = 270 nm and Em = 352 nm of NosL and its mutants showing formation of aldehyde **108** at 24.3 min. The aldehyde **108** was detected in all variants except NosL R323K. B) LC-MS analysis showing EIC $[M-H]^-$ = 158 Da confirming formation of **108** in NosL WT. C) MS of aldehyde **108** generated with NosL WT.

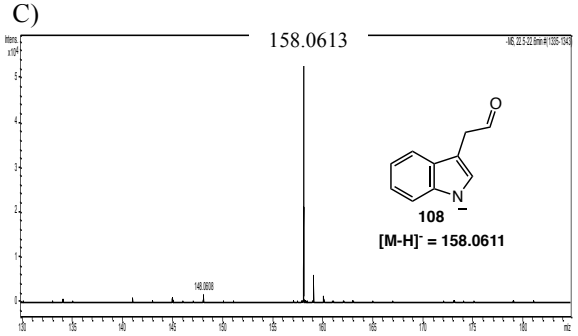


Figure 4.18 continued.

4.4.8 Reaction of NosL with 2'-chloro-L-tryptophan (119) and 2'-bromo-L-tryptophan (121)

2'-Halo-L-tryptophan (Br, Cl) was designed to trap a radical on C_3' by exploiting the very rapid cleavage of a C-halogen bond beta to a radical center.[84-85] When the NosL reaction was carried out with 2'-chloro-L-tryptophan (**119**), HPLC analysis showed complete substrate consumption and exclusive formation of a new product was observed (Figure 4.19). This product was identified as 3-methylindole-2-carboxylic acid (**15**) by LC-MS and chromatographic identity to an authentic standard. Similar results were obtained with 2'-bromo-L-tryptophan (**121**).

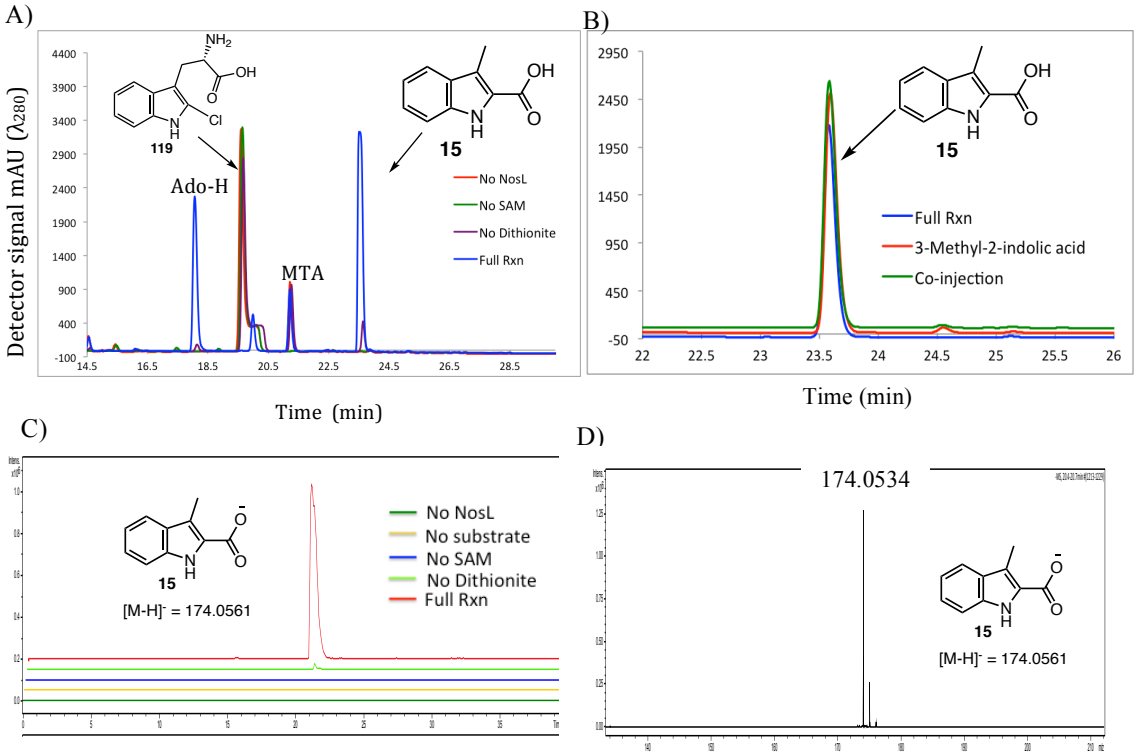


Figure 4.19: Analysis of the NosL reaction with 2'-chloro-L-tryptophan (**119**). A) HPLC analysis at 280 nm shows complete consumption of substrate **119** and formation of **15** in the 'Full Rxn' sample. B) Co-injection of 'Full Rxn' sample with authentic standard of **15** confirming its identity. C) LC-MS analysis showing EIC $[M-H]^-$ = 174 Da confirming formation of product **15**. E) MS of product **15**.

When 2'-chloro-L-tryptophan (**119**) was used as a substrate for the mutants of NosL, 3-methyl-2-indolic acid (**15**) was observed as the exclusive product with NosL Y90A, Y90F, C200A, S340A variants (Figure 4.20). Additionally, analogous to the reaction with L-tryptophan (**13**), NosL R323K variant generated the corresponding ketoacid **120** with **119** as a substrate.

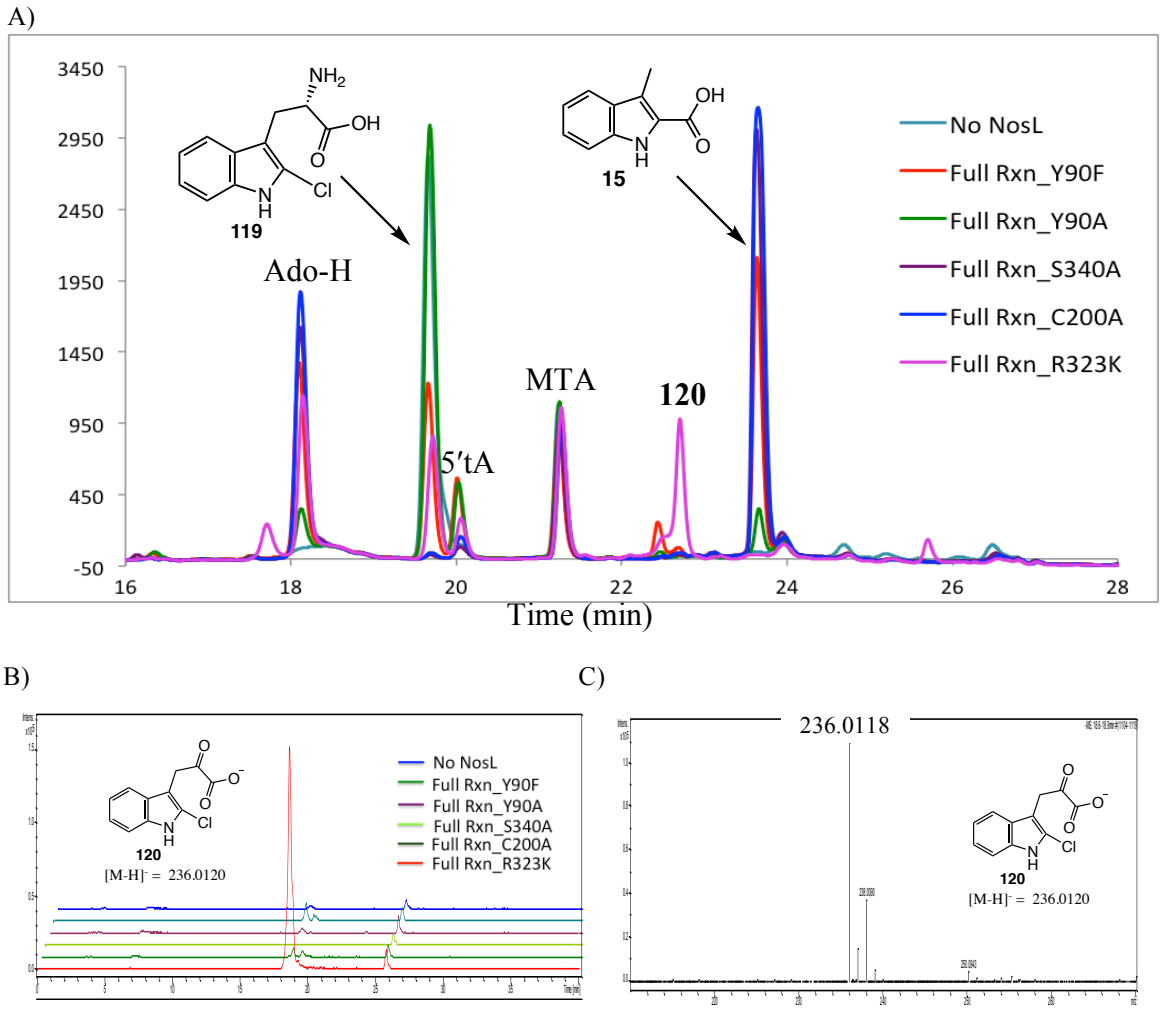


Figure 4.20: Analysis of the NosL variants with 2'-chloro-L-tryptophan (**119**). A) HPLC analysis at 280 nm of the reaction of NosL mutants with **119** showing formation of **15** in all variants except NosL R323K. B) LC-MS analysis showing EIC $[M-H]^-$ = 236.0 Da suggesting that the new product is **120** in NosL R323K variant. C) MS of product **120**.

During NosL catalyzed reaction with 2'-bromo-L-tryptophan (**121**), complete conversion to 3-methyl-2-indolic acid (**15**) was observed (Figure 4.21). Further studies of NosL mutants with **121** showed identical behavior as observed with **119** i.e. NosL Y90A, Y90F, C200A, S340A variants generated 3-methylindole-2-carboxylic acid (**15**) while NosL R323K formed the corresponding ketoacid.

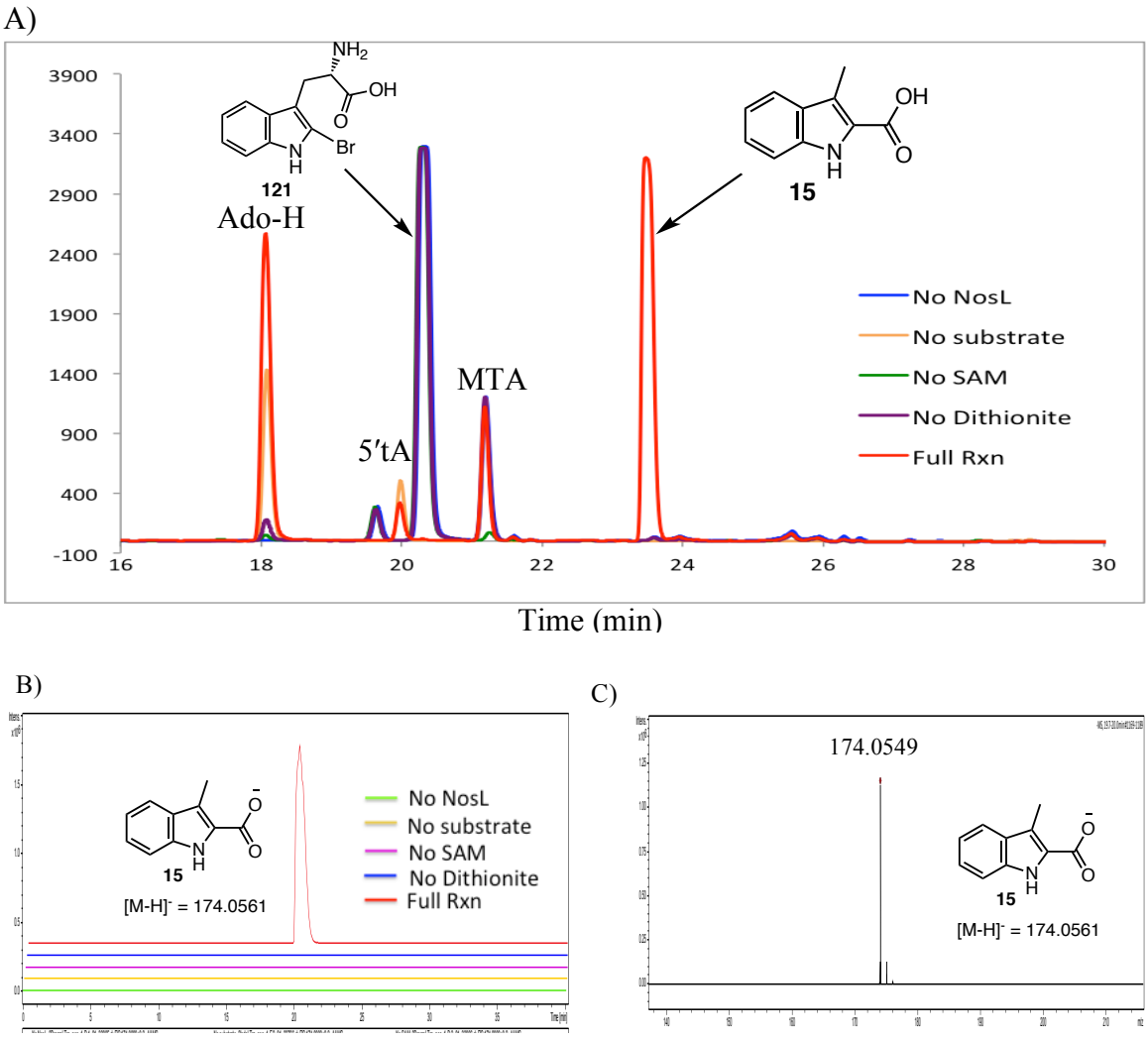


Figure 4.21: Analysis of the NosL reaction with 2'-bromo-L-tryptophan (**121**). A) HPLC analysis at 280 nm showing complete consumption of **121** in 'Full Rxn' sample, concomitant with formation of **15**. B) LC-MS analysis showing EIC $[M-H]^-$ = 154 Da confirming formation of **15**. C) MS of product **15**.

The enzymatic reaction mixture was next analyzed for the products derived from C_2 and the C_2 amino group of 2'-chloro-L-tryptophan (**119**). Formaldehyde, carbon monoxide and cyanide were each considered as possible products. The enzymatic reaction, treated with purpald or dansylhydrazine failed to yield the corresponding

formaldehyde adducts, thus excluding formaldehyde as a byproduct. Reduced hemoglobin also failed to detect carbon monoxide.[86] To test for cyanide (**82**) formation, the enzymatic reaction was treated with orthonapthalene dicarboxaldehyde (**122**) and taurine (**123**) to convert cyanide to the fluorescent derivative **124** (Figure 4.22).[87] HPLC analysis using fluorescence (Ex = 418 nm, Em = 454 nm) showed a new product confirmed as **124** by LC-MS and chromatographic identity with an authentic standard. The ratio of cyanide (**82**) to 3-methylindole-2-carboxylic acid (**15**) was ~ 0.9:1.

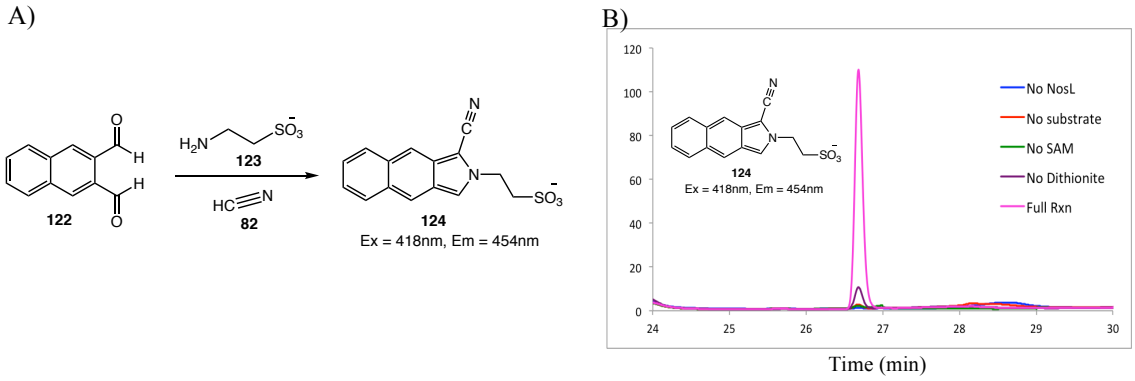


Figure 4.22: Detection of cyanide (**82**) as a byproduct of the NosL catalyzed reaction with 2'-chloro-L-tryptophan (**119**). A) Reaction used to convert cyanide to a fluorescent product **124**. B) HPLC analysis of the NosL/2'-chloro-L-tryptophan (**119**) reaction mixture after derivatization as shown in panel A. C) LC-MS analysis after derivatization showing EIC $[M-H]^-$ = 299 Da suggests formation of **124**. D) MS of the derivative **124**. E) LC-MS analysis showing co-injection of derivatized 'Full Rxn' sample with authentic standard of **124** generated using cyanide (**82**).

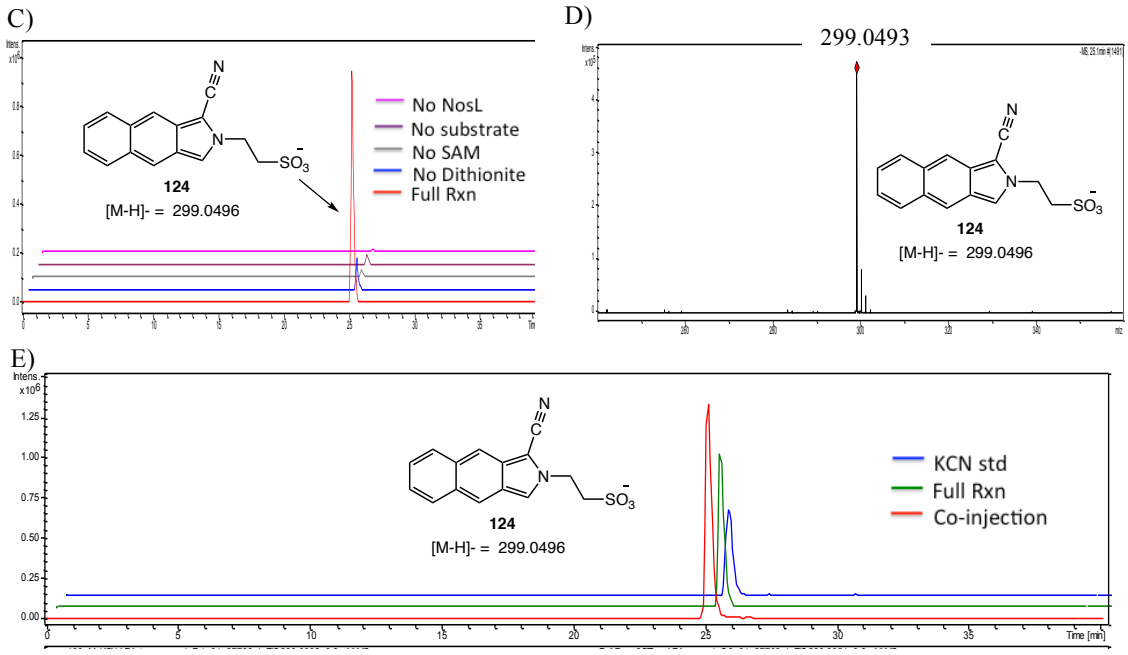


Figure 4.22 continued.

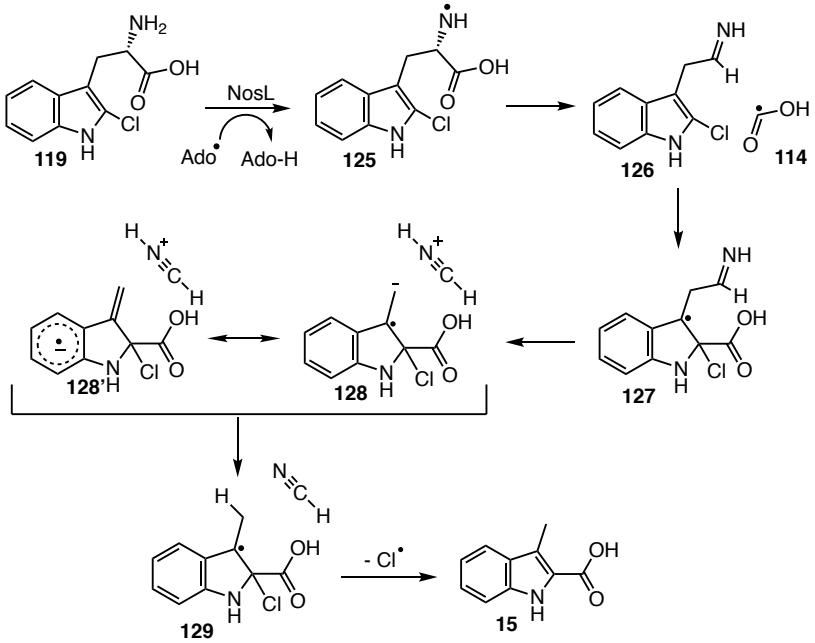


Figure 4.23: Mechanistic proposal of the NosL catalyzed reaction with 2'-chloro-L-tryptophan (**119**).

A mechanistic proposal for the NosL catalyzed reaction with 2'-chloro-L-tryptophan (**119**) is shown in Figure 4.23. Hydrogen atom abstraction from the amino group of 2'-chloro-L-tryptophan gives **125**. A β-scission reaction releases the formyl radical (**114**), which then adds to C_2' of the indole to give **127**. Loss of H_2CN^+ from **127** gives the stabilized radical anion **128** which can also be viewed as an aryl radical anion **128'**. Protonation at C_3, potentially using the strongly acidic H_2CN^+, followed by loss of a chlorine atom completes the reaction.

4.4.9 Reaction of NosL with 2'-methyl-L-tryptophan (130)

When the reaction of NosL was performed with 2'-methyl-L-tryptophan (**130**), formation of 2,3-dimethyl indole (**131**) and cyanide (**82**) was observed (Figure 4.24).

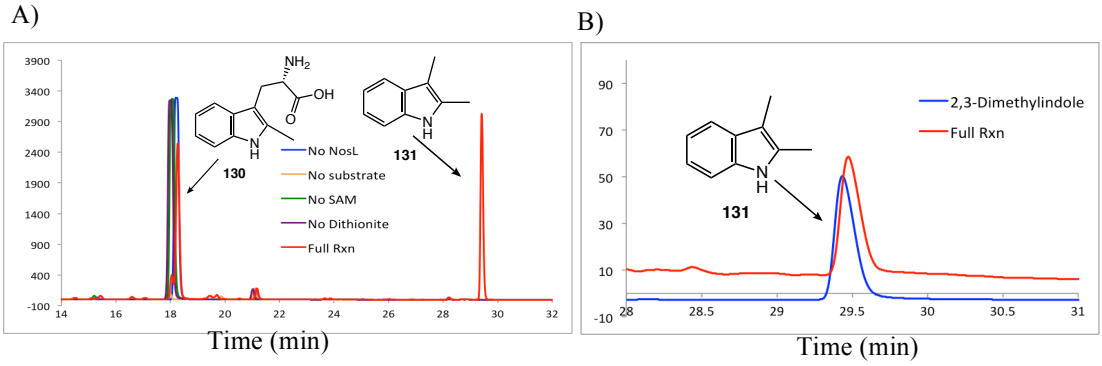


Figure 4.24: HPLC analysis of the NosL reaction with 2'-methyl-L-tryptophan (**130**). A) HPLC analysis at 288 nm shows consumption of substrate and formation of a new signal at 29.5 min. B) HPLC analysis at 288 nm shows co-migration of the authentic standard of 2,3-dimethyl indole (**131**) and the compound generated during NosL reaction with **130**. C) Fluorescence HPLC analysis at Ex = 418 nm and Em = 454 nm confirms formation of cyanide (**82**) in 'Full Rxn' sample by trapping **124**.

C)

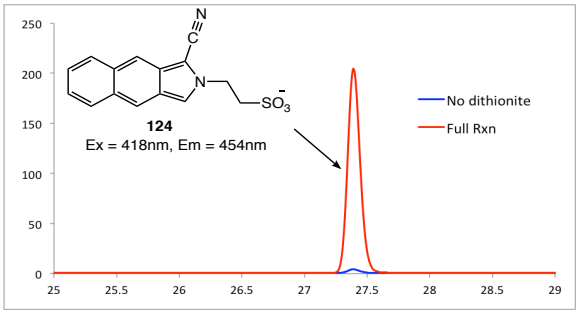


Figure 4.24 continued.

4.4.10 Analysis of the NosL reaction using ^{13}C and ^{15}N isotopologues of L-tryptophan

Next we turned our attention to the identification of the products of dithionite-reduced NosL reaction with L-tryptophan (**13**). It was reported that 3-methylindole (**92**) is the major heterocyclic product of this reaction and glyoxal (**101**), formaldehyde (**87**) and ammonia (**88**) were assumed to be the C_1/C_2, the C_2 and C_2 amino-derived products respectively, based on the detection, at trace levels by LC-MS, of glyoxal (**101**) and formaldehyde (**87**) in the reaction mixture (Figure 4.2D). The quantitative detection of cyanide (**82**) as a NosL reaction product in the 2'-chloro-L-tryptophan (**119**) and 2'-methyl-L-tryptophan (**130)** reactions cast doubt on this assumption.

Isotopologues of L-tryptophan (**13**) were synthesized chemo-enzymatically by feeding indole and isotopically labeled L-serine to tryptophan synthase in presence of PLP (Figure 4.25).[88] Isotopologues [1,2,3-$^{13}C_3$, 2-amino-$^{15}N_1$]-L-tryptophan (**13d**) and [1-$^{13}C_1$]-L-tryptophan (**13e**) were synthesized using tryptophan synthase while [$^{15}N_2$]-L-tryptophan (**13f**) was obtained commercially.

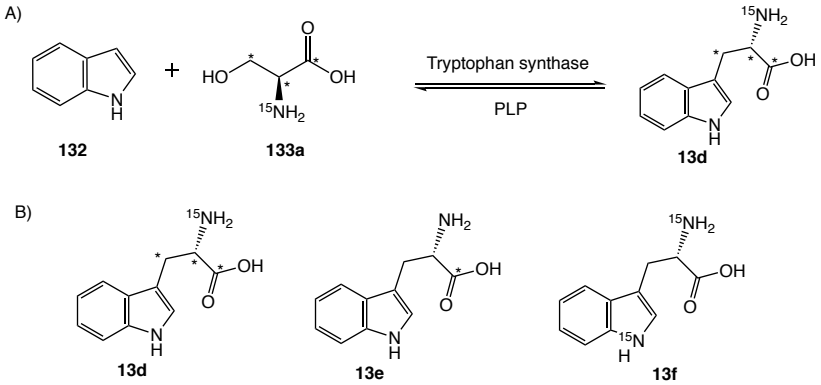


Figure 4.25: Isotopologues of L-tryptophan (**13**). A) Scheme depicting chemo-enzymatic synthesis of the isotopologues using tryptophan synthase. B) Isotopologues used to investigate products of the NosL catalyzed reaction.

When the NosL reaction was performed with [1,2,3-$^{13}C_3$, 2-amino-$^{15}N_1$]-L-tryptophan (**13d**), ^{13}C-NMR analysis of the reaction mixture revealed only two new signals (8 ppm and 160 ppm, Figure 4.26A). The 8 ppm signal was assigned to the methyl group of 3-methylindole (**92**) and the 160 ppm signal was assigned to cyanide (**82**) and/or bicarbonate by spiking with authentic standards and recording the spectra.

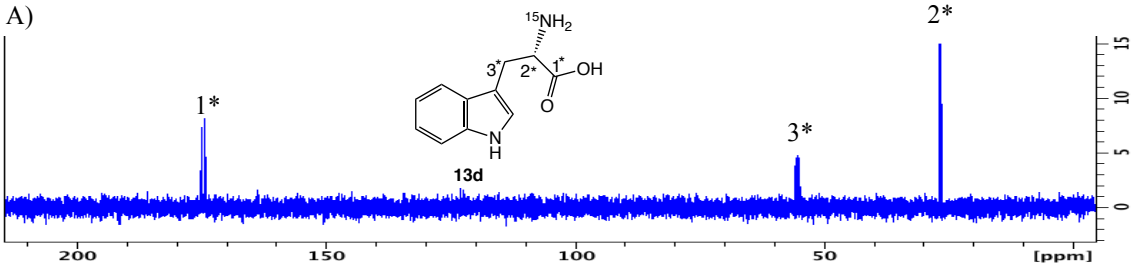


Figure 4.26: ^{13}C-NMR analysis of NosL reaction with [1,2,3-$^{13}C_3$, 2-amino-$^{15}N_1$]-L-tryptophan (**13d**). A) NMR spectra of the substrate **13d**. B) Full reaction sample shows two signals at 8 ppm and 160 ppm. C) Full reaction sample spiked with 20 mM KCN (final conc.) shows a single peak at 160 ppm. Magnified region at 160 ppm is shown in the small panel. D) Full reaction sample spiked with 50 mM $NaHCO_3$ (final conc.) shows a single peak at 160 ppm. Region at 160 ppm is magnified in the small panel.

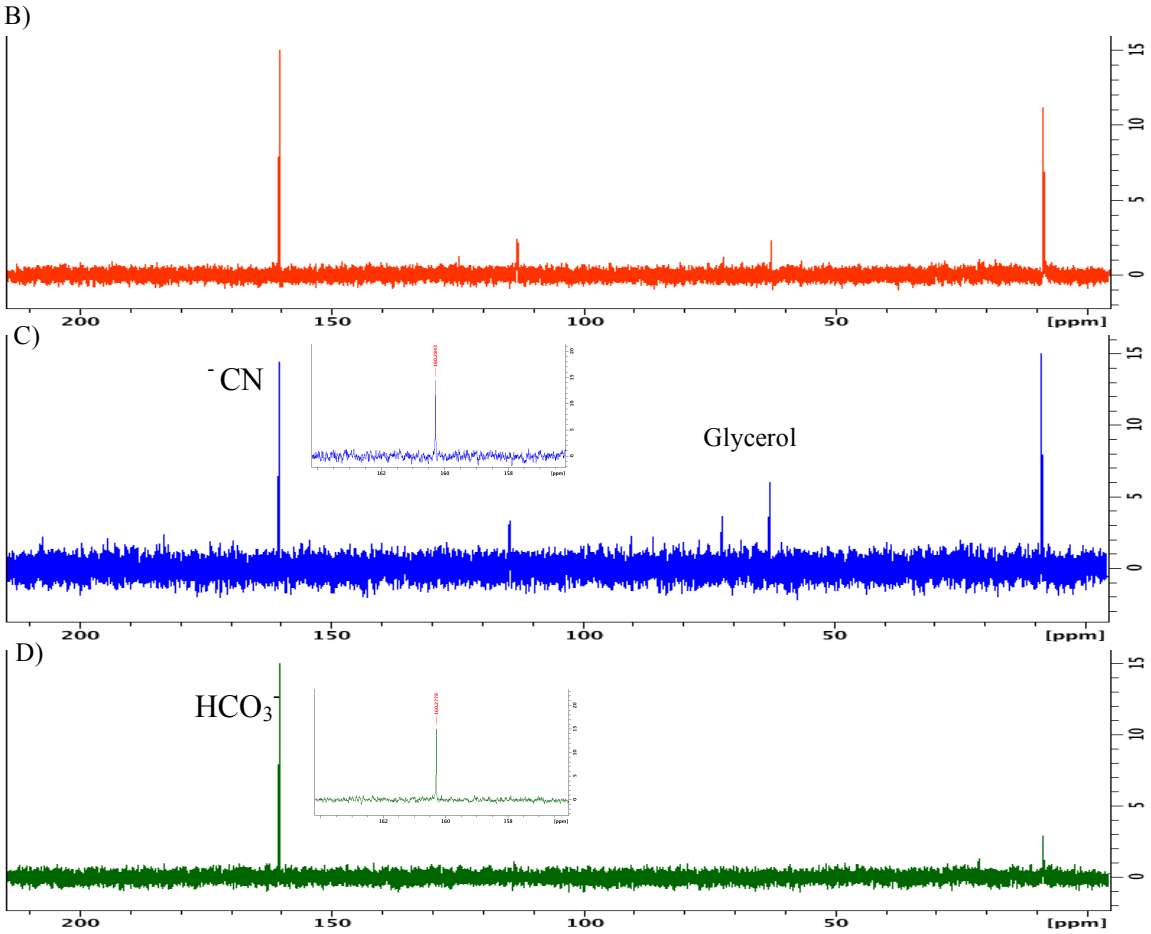


Figure 4.26 continued.

Next, ^{13}C-NMR spectra were recorded of the NosL catalyzed reaction with [1-$^{13}C_1$]-L-tryptophan (**13e**). The sample was further spiked with $NaHCO_3$ and a single signal was observed at 160 ppm (Figure 4.27). This confirms that the C_1 fragment of L-tryptophan (**13**) is lost as CO_2, which is trapped as bicarbonate under buffer conditions.

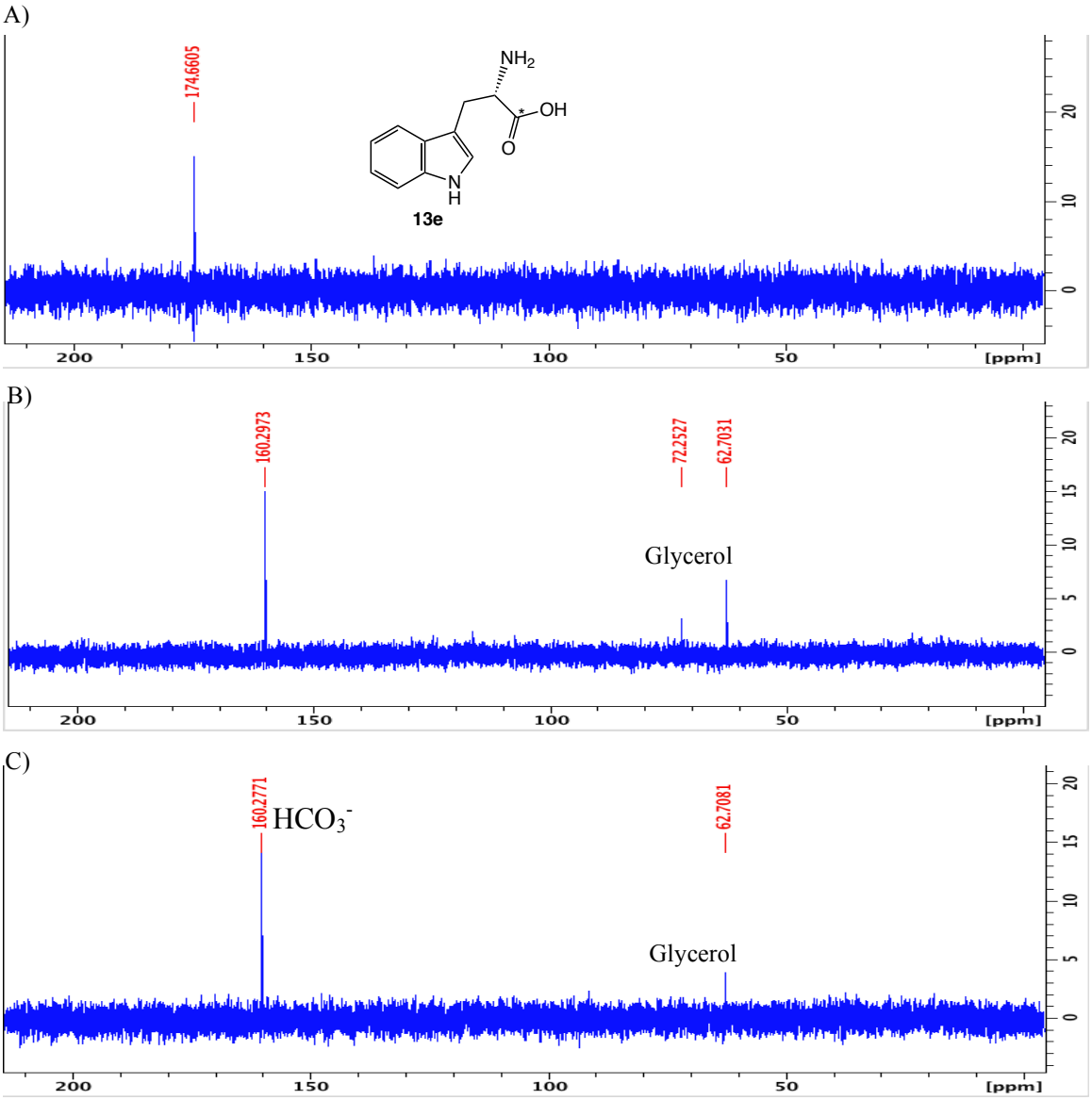


Figure 4.27: ^{13}C-NMR analysis of NosL reaction with [1-$^{13}C_1$]-L-tryptophan (**13e**). A) NMR spectra of the substrate **13e**. B) Full reaction sample shows formation of a unique signal at 160 ppm. C) Full reaction sample was spiked with 50 mM $NaHCO_3$. It shows a unique signal at 160 ppm confirming the identity of the C_1 fragment.

To test formation of CO_2/bicarbonate (**134**), we treated the NosL enzymatic reaction with phosphoenolpyruvate carboxylase (PEPC) in presence of phosphoenolpyruvate (**135**). Any bicarbonate present in solution will be derivatized by phosphoenolpyruvate (**135**) and generate oxaloacetate (**136**) as shown in Figure 4.28. To establish the identity of oxaloacetate (**136**), the reaction mixture was derivatized with o-(2,3,4,5,6-pentafluorobenzyl)hydroxylamine.HCl (**137**, PFBHA) and analyzed on LC-MS.[89]

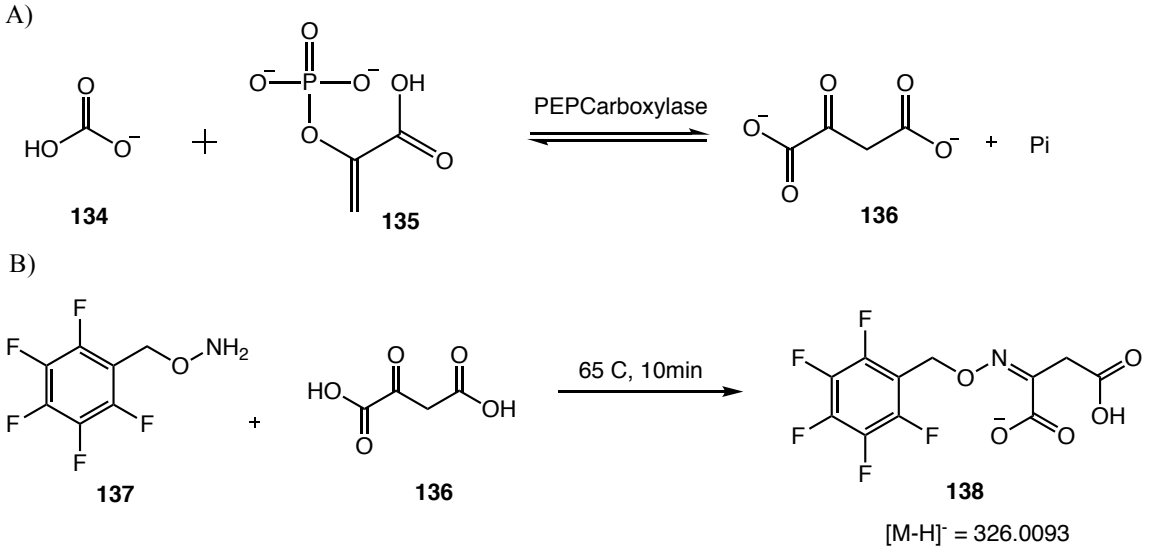


Figure 4.28: Derivatization of bicarbonate using PEPCarboxylase. A) Scheme of PEPC catalyzed reaction, showing formation of oxaloacetate (**136**). B) Derivatization of oxaloacetate (**136**) with PFBHA (**137**) for LC-MS detection.

Reaction of NosL with [1-$^{13}C_1$]-L-tryptophan (**13e**) followed by PEPC treatment suggests formation of CO_2/bicarbonate from C_1 fragment of L-tryptophan (Figure 4.29). Thus PEPC can be used to derivatize all the CO_2/bicarbonate generated.

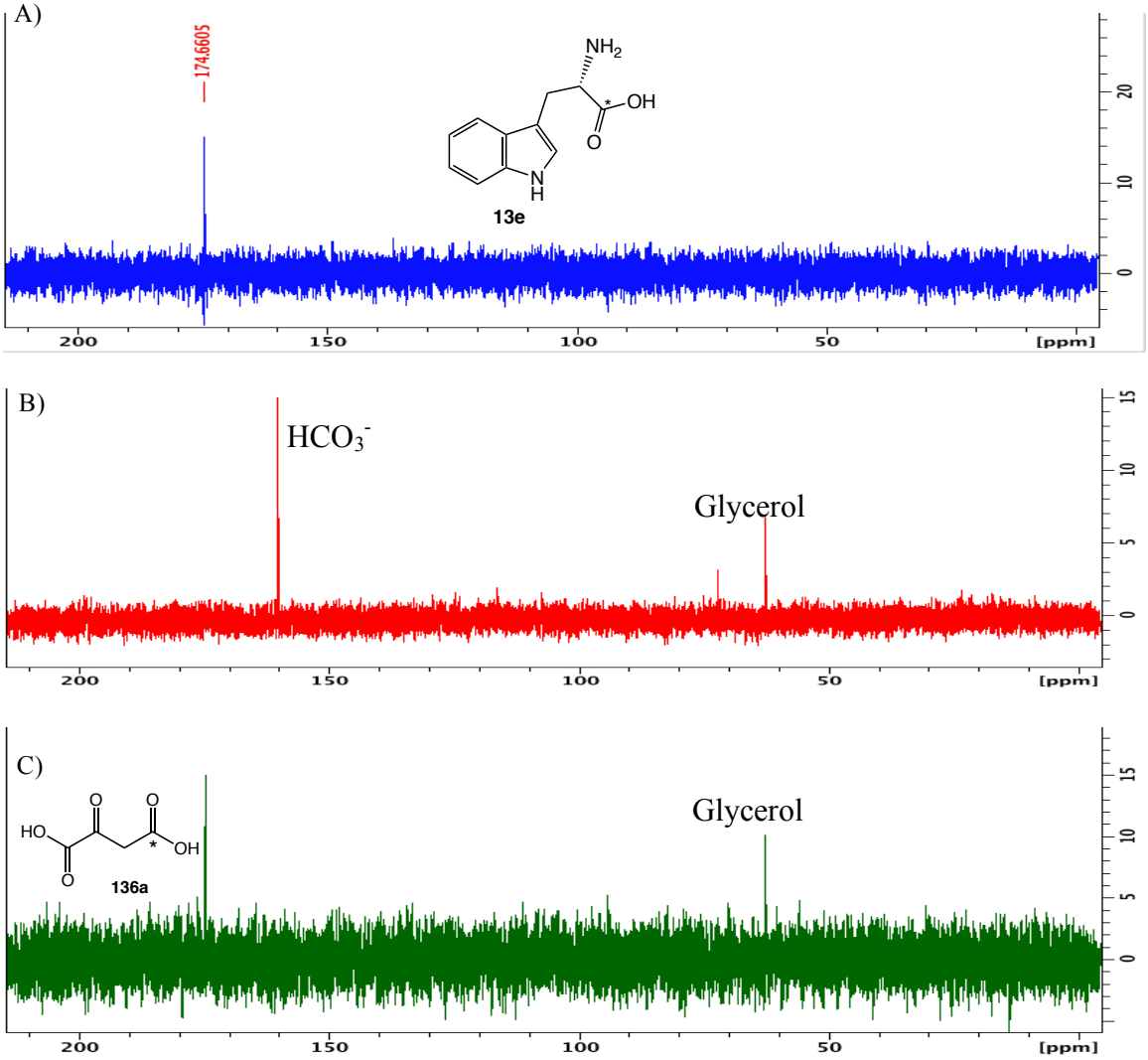


Figure 4.29: ^{13}C-NMR analysis of NosL reaction with [1-$^{13}C_1$]-L-tryptophan (**13e**). A) NMR of the substrate **13e**. B) NMR of the Full reaction sample when **13e** is used as a substrate shows formation of a single peak at 160 ppm. C) When the full reaction sample is treated with PEPC, the signal at 160 ppm disappears and a new signal at 175 ppm is observed indicating formation of oxaloacetate (**136**).

Reaction of NosL with [1,2,3-$^{13}C_3$, 2-amino-$^{15}N_1$]-L-tryptophan (**13d**) followed by treatment with PEPC suggests formation of CO_2/bicarbonate from the C_1 fragment and CN from the C_2 fragment of L-tryptophan (**13**).

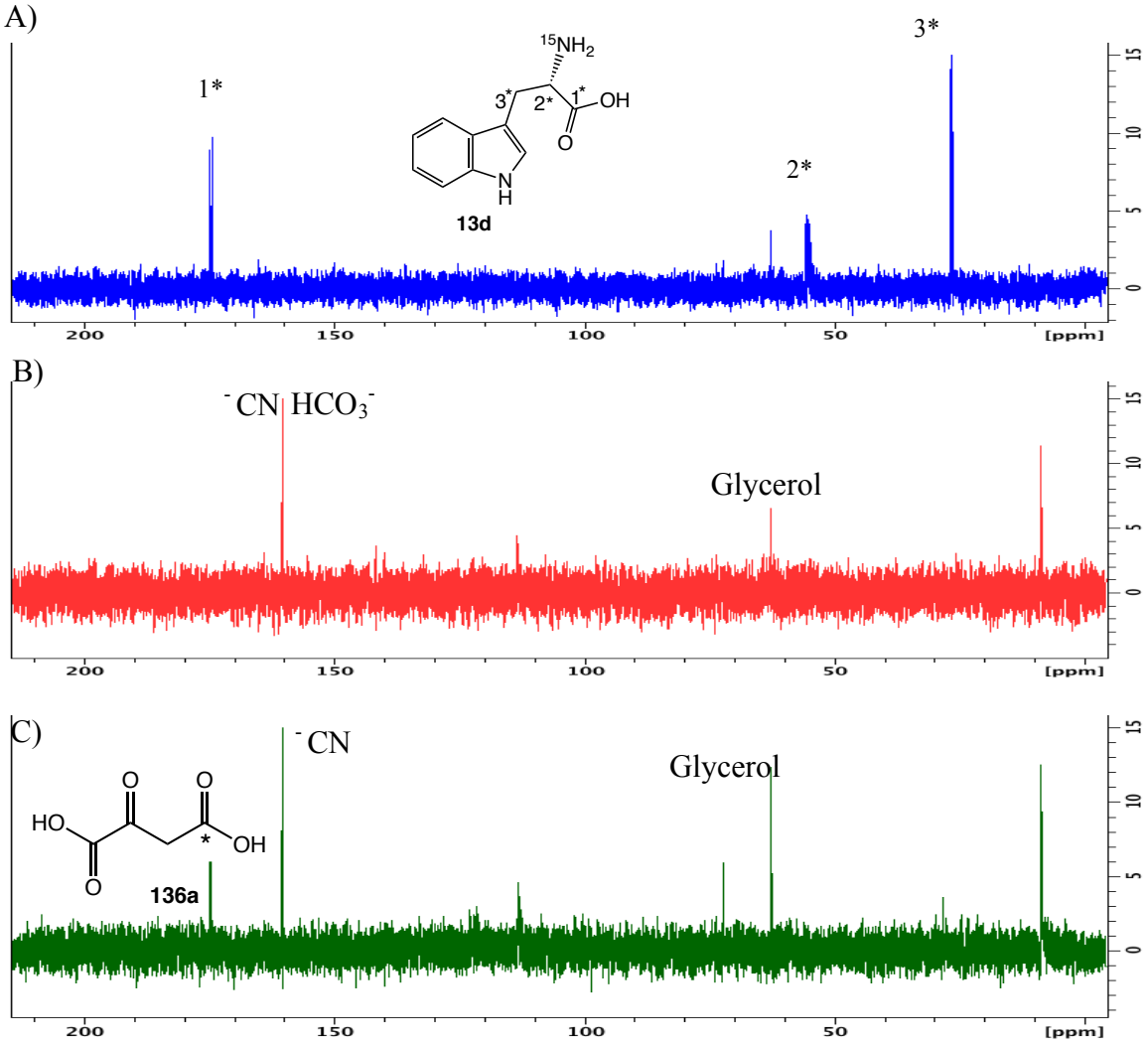


Figure 4.30: ^{13}C-NMR analysis of NosL reaction with [1,2,3-$^{13}C_3$, 2-amino-$^{15}N_1$]-L-tryptophan (**13d**). A) NMR of the substrate [1,2,3-$^{13}C_3$, 2-amino-$^{15}N_1$]-L-tryptophan (**13d**). B) NMR of the Full reaction sample when **13d** is used as a substrate shows formation of a single peak at 160 ppm. C) When the full reaction sample is treated with PEPC, a new signal appears at 175 ppm indicates formation of oxaloacetate while signal at 160 ppm indicates ^{-}CN formation.

The bicarbonate that is converted to oxaloacetate using PEPC was further derivatized with PFBHA (**137**) for LC-MS analysis of adduct (**138**). Figure 4.31 shows MS spectra of **138** generated when unlabeled L-tryptophan (**13**), [1,2,3-$^{13}C_3$, 2-amino-$^{15}N_1$]-L-tryptophan (**13d**) and [1-$^{13}C_1$]-L-tryptophan (**13e**) were used as substrates.

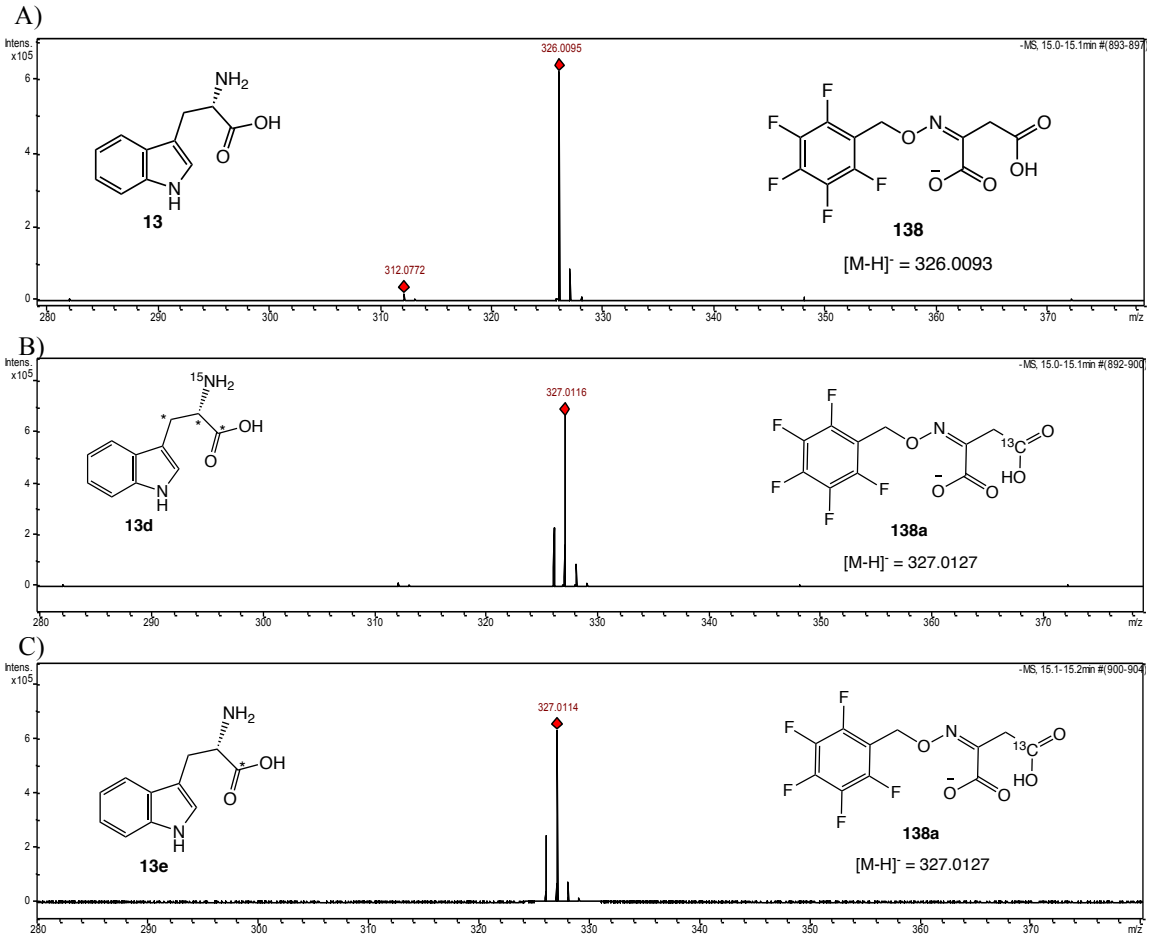


Figure 4.31: LC-MS analysis of oxaloacetate generated by PEPC which is further derivatized with PFBHA (**137**) when different isotopologues of L-tryptophan are used as substrates for NosL reaction. A) MS of derivatized compound **138** is $[M-H]^-$ = 326.0 Da when unlabeled L-tryptophan (**13**) is used as substrate. B) MS of derivatized compound **138a** is $[M-H]^-$ = 327.0 Da when [1,2,3-$^{13}C_3$, 2-amino-$^{15}N_1$]-L-tryptophan (**13d**) is used as substrate. C) MS of derivatized compound **138a** is $[M-H]^-$ = 327.0 Da when [1-$^{13}C_1$]-L-tryptophan (**13e**) is used as substrate.

To further confirm the formation of cyanide (**82**), its derivatization was carried out using **122** and **123** as shown in Figure 4.22A. MS of the derivatized product **124** was analyzed to investigate the incorporation of C_2 and C_2 amino fragments from isotopologues of L-tryptophan as shown in Figure 4.32.

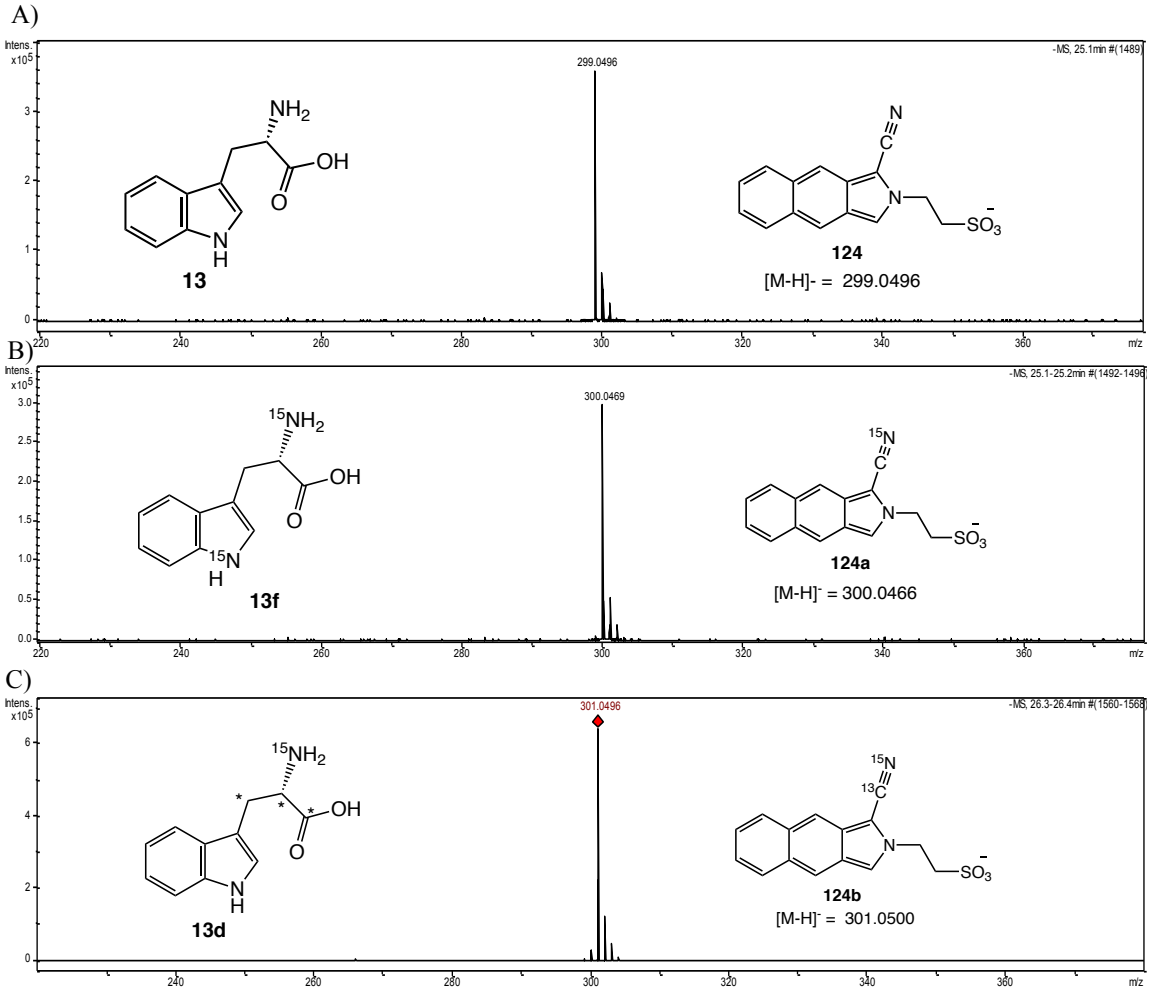


Figure 4.32: LC-MS analysis of derivatized cyanide when different isotopologues of L-tryptophan (**13**) are used as substrates for NosL reaction. A) MS of derivatized compound **124** is $[M-H]^-$ = 299 Da when unlabeled L-tryptophan (**13**) is used as substrate. B) MS of derivatized compound **124a** is $[M-H]^-$ = 300 Da when [$^{15}N_2$]-L-tryptophan (**13f**) is used as substrate. C) MS of derivatized compound **124b** is $[M-H]^-$ = 301 Da when [1,2,3-$^{13}C_3$, 2-amino-$^{15}N_1$]-L-tryptophan (**13d**) is used as substrate.

After cyanide (**82**) was observed as a product of the native NosL catalyzed reaction with L-tryptophan (**13**), we investigated its formation in the previously described variants of NosL. Derivatization of cyanide was carried out using **122** and **123** as shown in Figure 4.22A for the different variants of NosL. As shown in Figure 4.33, the fluorescence (Ex = 418 nm, Em = 454 nm) of the derivatized product **124** was observed on HPLC for NosL WT, Y90F, Y90A, C200A and S340A. However, cyanide was not observed with NosL R323K. The ratio of 3-methylindole (**92**) to cyanide (**82**) in NosL WT and its variants is close to 1:0.9. This demonstrates that cyanide (**82**) is the major byproduct of the NosL catalyzed reaction with L-tryptophan (**13**).

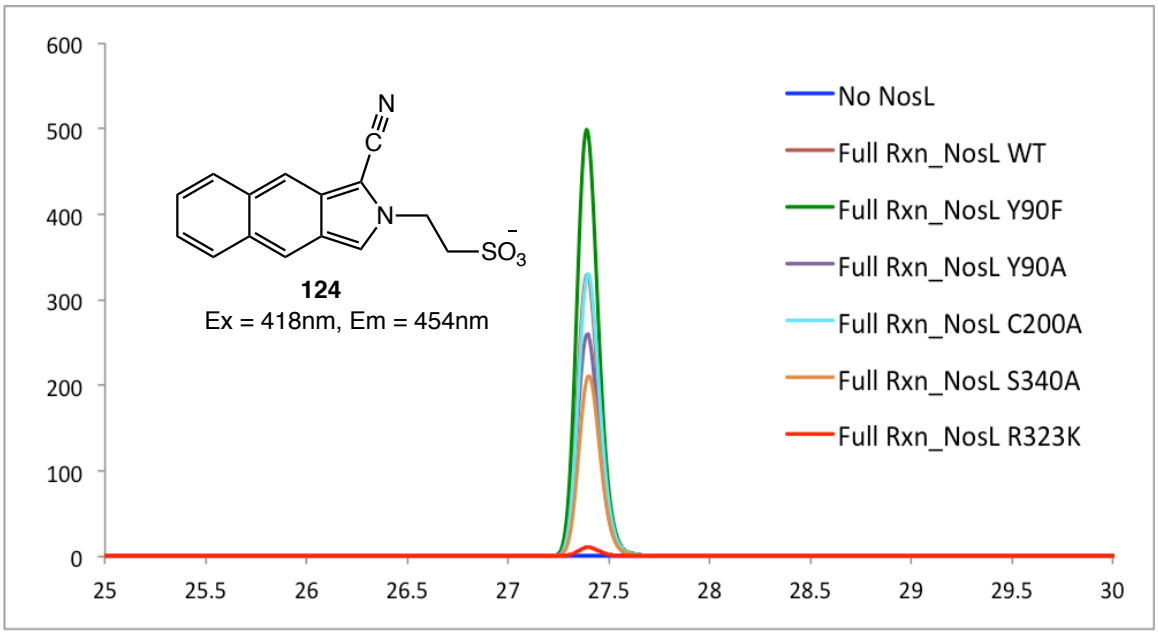


Figure 4.33: Fluorescence detection of cyanide (**82**) formed during the reaction of NosL variants with L-tryptophan (**13**). Quantitative cyanide was detected in all variants tested except R323K.

4.4.11 Revised mechanistic proposal of the NosL catalyzed reaction

The results mentioned in this section can be integrated into the mechanistic proposal for the NosL-catalyzed reaction with L-tryptophan (**1**) as shown in Figure 4.34. In this revised mechanism, the initially formed amine radical (**14**) undergoes a β-scission reaction to form **107** and the formyl radical (**114**) which then adds to the C_2' of the indole to form **139**. Loss of H_2CN^+ by a second β-scission reaction gives **140** which can also be described as an aryl radical anion **140'**. Protonation of this radical by the strongly acidic H_2CN^+ followed by electron transfer back to the $[4Fe\text{-}4S]^{+2}$ cluster gives the cation **142**. This can either lose CO_2 or a proton to give the experimentally observed reaction products **15** and **92** respectively. Alternatively loss of a formyl radical from **141** could also generate **92**.

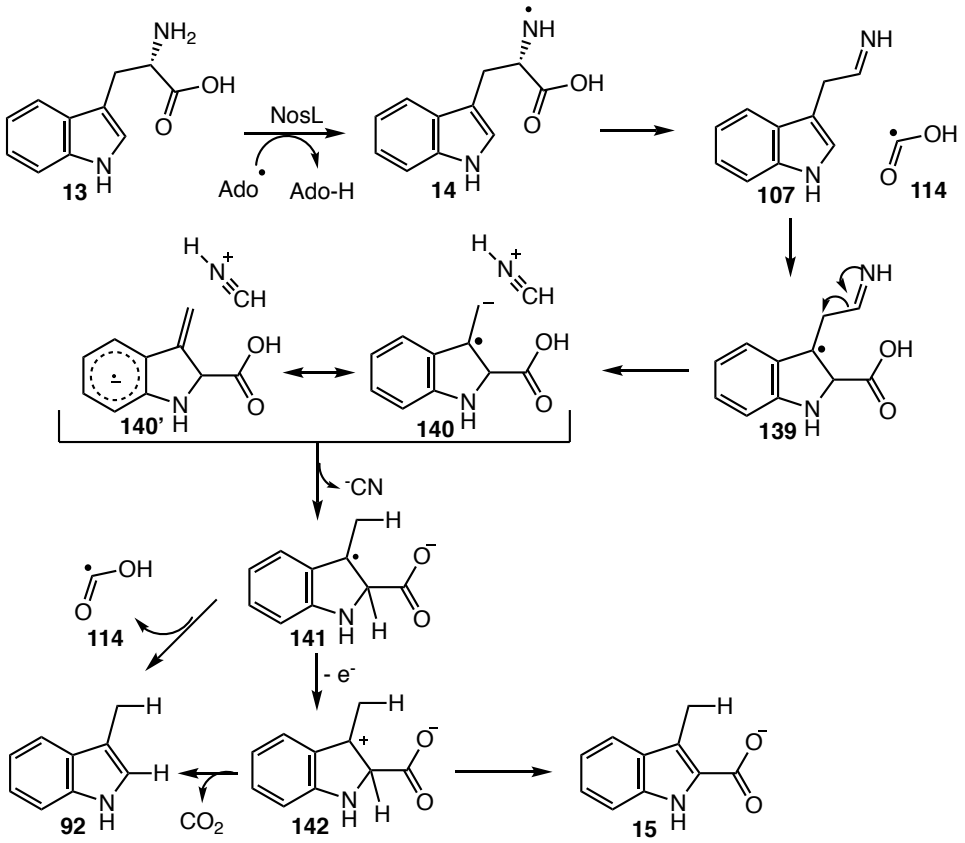


Figure 4.34: Revised mechanistic proposal of the NosL catalyzed reaction integrating all the data mentioned in this section.

4.5 Conclusion

In this section, we described a set of experiments using substrate analogs, mutants and isotopes of L-tryptophan (**13**) to probe the mechanism of NosL. We provide biochemical evidence in support of the structural studies showing that the 5′-deoxyadenosyl radical (**6**, Ado$^{\bullet}$) abstracts a hydrogen atom from the amino group of L-tryptophan (**13**). We further demonstrated that D-tryptophan (**104**) is a substrate for NosL but shows relaxed regio control of the first β-scission reaction. Mutagenesis studies

confirm that Arg323 is important for controlling the regiochemistry of the β-scission reaction. Based on the early studies, a mechanism involving C_α-C_β bond scission after the H-atom abstraction from L-tryptophan (**13**) is proposed.

With L-tryptophan thiocarboxylate (**115**), we provide evidence that the C_α-C(O) cleavage is the next step after the H-atom abstraction. Detection and characterization of aldehyde **108** in NosL WT and its variants also supports the C_α-C(O) cleavage instead of previously proposed β-bond scission involving the C_α-C_β bond. Use of 2'-chloro-L-tryptophan (**119**) or 2'-bromo-L-tryptophan (**121**) as substrates lead to the exclusive formation of 3-methylindole-2-carboxylic acid (**15**) and cyanide (**82**). Further investigation revealed that cyanide (**82**) is also a byproduct of the NosL catalyzed reaction with L-tryptophan (**13**).

The revised mechanistic proposal is consistent with the EPR spectroscopy that provides evidence for the formation of **139**.[90] It is also consistent with the NosL reaction with L-tryptophan thiocarboxylate (**115**), which blocks the formation of **139** and demonstrates that the formation of 3-methylindole (**92**) can only occur after addition of the formyl radical to the indole. Finally, it is consistent with the exclusive formation of 3-methylindole-2-carboxylic acid (**15**) [no 3-methylindole (**92**) detected] as the sole product when 2'-chloro-L-tryptophan (**119**) is used as the substrate. The identification of cyanide (**82**) as a major NosL byproduct has important mechanistic implications and connects NosL chemistry with that of HydG – the radical SAM enzyme that forms cyanide (**82**) and carbon monoxide (**83**) from tyrosine during the biosynthesis of the metallo cluster of the [FeFe] hydrogenase (Figure 4.2B).[91-93]

4.6 Experimental procedures

Materials: All chemicals were purchased from Sigma-Aldrich (now MilliporeSigma) unless specified otherwise. LB medium (Lennox formulation) was from EMD Millipore. Kanamycin was from Teknova and IPTG was obtained from Lab scientific Inc. Chloramphenicol was from Fisher Scientific. HPLC and LC-MS solvents were purchased from EMD and were used without further purification. Histrap column (5 ml) was obtained from GE healthcare. Econo-Pack 10DG and Bio-spin 6 desalting columns were purchased from Bio-Rad Laboratories. D_2O, [$^{13}C_3$-$^{15}N_1$]-L-serine, [1-$^{13}C_1$]-L-serine and $^{15}N_2$-L-tryptophan were purchased from Cambridge Isotope Laboratories Inc. Large cultures were grown and overexpressed in 2.5 L baffled ultra yield flasks and sealed with AirOtop seals, both of which were from Thomson Instrument Company. *E.coli* containing plasmid of tryptophan synthase (pSTB7) was ordered from ATCC. Phosphoenolpyruvate carboxylase (PEPC) was purchased from ICN Biomedicals Inc. (now MP Biomedicals). Other enzymes like L-amino acid oxidase from *Crotalus adamanteus*, D-amino acid oxidase, α-chymotrypsin and amano acylase were obtained from Sigma-Aldrich. 2'-methyl-D,L-tryptophan (**130a**) was obtained from Chem-Impex International Inc.

Overexpression and purification of NosL and its mutants: Gene of *nosl* from *Streptomyces actuosus* was cloned into pTHT vector (derivative of pET28b vector with TEV protease cleavage site after N-terminal His-tag). NosL was co-expressed with a plasmid encoding the *suf* operon in *E.coli* BL21(DE3) for *in vivo* assembly of the [4Fe-4S] cluster. A starter culture was grown overnight in LB media containing kanamycin

(40 μg/ml) and chloramphenicol (25 μg/ml). 60 ml of this culture was added to 6L LB media (4x1.5L flasks) with antibiotics. The flasks were sealed with AirOtop and grown at 37°C with shaking (180 rpm) till OD_{600}~0.6. The flasks were then incubated at 4 °C for ~1 hr without shaking. Each flask (1.5L) was supplemented with 120 mg ferrous ammonium sulfate and 120 mg cysteine. Cultures were then induced with 200 μM IPTG followed by incubation at 15 °C for ~16 hr with shaking (110 rpm). The cells were harvested by centrifugation and stored in liquid nitrogen until further use. Typical yields were 13-15 gm of cell pellet (wet weight) from 6L cell culture.

All steps for NosL purification were carried out in an anaerobic chamber (COY laboratories). Cell pellets were thawed and resuspended in 60-70 ml of lysis buffer (100 mM Tris-HCl, pH 7.5) at room temperature in the presence of lysozyme (8-10 mg) and benzonase (1000 units). The suspension was stirred for ~1.5 hr on an ice-bath and further sonicated to lyse the cells. Cell debris was removed by centrifugation and the lysate was loaded onto a Histrap column pre-equilibrated with lysis buffer. The column was washed with 10 column volumes of wash buffer (100 mM Tris-HCl, 30 mM imidazole, 300 mM NaCl, pH 7.5). The protein was then eluted from the histrap column with elution buffer (100 mM Tris-HCl, 250 mM imidazole, 300 mM NaCl, pH 7.5). The dark colored fractions were pooled and buffer exchanged to 50 mM potassium phosphate, 10% glycerol, pH7.5 using an Econo-Pac 10DG desalting column. The purified enzyme was stored submerged in liquid nitrogen until further use.

Protein concentration was measured by the absorbance at 280 nm (A_{280}) with an extinction coefficient calculated using the ProtParam tool of the ExPASy proteomics

server. Typical yields were ~35 mg of enzyme per liter of cell culture. Iron content was determined using ferene assay while sulfide content was determined using methylene blue assay (methods previously described). Iron content of the purified enzyme was 4.2±0.2 and the sulfide content was 3.7±0.1. All mutants of NosL were also purified using the above-described protocol.

Overexpression and purification of FlavodoxinA/Flavodoxin reductase: The following protocol was used for both *E.coli* flavodoxin A (FldA) and flavodoxin reductase (FldR). Respective plasmids were transformed into *E.coli* BL21(DE3) cells. A starter culture was grown overnight in LB media containing kanamycin (40 μg/ml). 30 ml of this overnight culture was added to 3L LB media (2x1.5L flasks) with kanamycin. The flasks were sealed with AirOtop and grown at 37 °C with shaking (210 rpm) till OD_{600}~0.6. The flasks were then incubated at 4 °C for ~30 min without shaking. Cultures were then induced with 500 μM IPTG followed by incubation at 30 °C for ~6 hr with shaking (180rpm). The cells were then harvested by centrifugation and stored at -80 °C until further use.

Protein purification was carried out under aerobic conditions. Cells were resuspended in 40 ml of lysis buffer (50 mM Tris-HCl, 10 mM imidazole, 150 mM NaCl, pH 7.5). This suspension was sonicated and centrifuged to separate debris. Cell lysate was loaded on a Histrap column pre-equilibrated with lysis buffer. The column was washed with 100 ml of washing buffer (50 mM Tris-HCl, 20 mM imidazole, 150 mM NaCl, pH 7.5) including 50 μM flavin (FMN for FldA and FAD for FldR). Elution

was carried out with elution buffer (50 mM Tris-HCl, 250 mM imidazole, 150 mM NaCl, pH 7.5). The protein fractions were concentrated and buffer exchanged with an Econo-Pac 10DG desalting column into 100 mM potassium phosphate, 150 mM NaCl, 10% glycerol, 1 mM DTT, pH 7.5 under anaerobic conditions. The purified proteins were stored at -80°C until further use.

Enzymatic reaction conditions for NosL: All NosL enzymatic reactions were carried out in an anaerobic chamber containing 95% nitrogen and 5% hydrogen. Typical enzymatic reaction was performed in 100 mM phosphate buffer, pH7.5 containing NosL (75-150 μM), dithionite (600 μM), L-tryptophan / substrate analog (400 μM - 500 μM) and SAM (600 μM).

When assays were performed in the presence of the flavodoxin system, NosL was desalted using Bio-spin 6 columns (Bio-rad) into 100 mM phosphate buffer, pH 7.5. The components of the enzymatic reaction were NosL (75-100 μM), FldA (60 μM), FldR (40 μM), NADPH (750 μM), L-tryptophan / substrate analog (500 μM) and SAM (600 μM). The enzymatic reactions were incubated at room temperature for 2-3 hr. The enzymes were removed either by ultrafiltration using 10 kda cut-off filters (VWR) or by 5-10% TFA (final conc.) and further analyzed by HPLC and LC-MS.

For NMR analysis of enzymatic reactions, NosL was desalted using a Bio-spin 6 column (Bio-Rad) into 100 mM phosphate buffer, pH 7.5 (to remove glycerol from the storage buffer). The components of the enzymatic reaction were NosL (75-100 μM), dithionite (750 μM), labeled L-tryptophan (400 μM) and SAM (600 μM). The enzymatic

reactions were transferred to NMR tubes and incubated in anaerobic chamber at room temperature for 2 hr.

Overexpression of tryptophan synthase and chemo-enzymatic synthesis of L-tryptophan isotopologues: Stock of *E.coli* BL21(DE3) cells containing plasmid of tryptophan synthase (pSTB7) was grown overnight in 20 ml LB media supplemented with 100 µg/ml of ampicillin. 15 ml of this overnight culture was used to inoculate 1.5L media containing 100 µg/ml of ampicillin. The cells were grown at 37 °C at 200 rpm for 24 hr. The cells were harvested by centrifuging at 9000 rpm for 10 min. The cells were divided in three falcon tubes with 3.5 gm wet cells in each falcon and stored at -80°C until further use.

To obtain the lysate, cells from one falcon tube were suspended in 35 ml of lysis buffer (50 mM KH_2PO_4, pH7.6). 2 mg of pyridoxal 5'-phosphate (PLP) was added to it and 4 cycles of sonication were performed (30 sec, 50% cycle, 65% power) at the interval of 10 min. The lysate was centrifuged at 18,000rpm for 20 min at 4 °C and filtered with 0.22 µm filters to remove cell debris. The lysate was stored in 4 °C for up to one month and was used whenever required in this period.

To synthesize the isotopologues of L-tryptophan (**13**), 10 mg of indole and 12 mg of respective isotopologues of L-serine were dissolved in 12 ml of lysis buffer. 0.5 mg of PLP was added to this solution followed by addition of 3 ml lysate. The reaction was incubated at 37 °C at 200 rpm for 3-4 hr. Proteins were removed by heat denaturing at 75 °C for 20 min followed by ultrafiltration using 10 kda cut-off

filters (VWR). Molecules were purified on HPLC using semi-prep column. Yield 5-8 mg product.

Derivatization of Cyanide:[87] A 1 ml stock solution for cyanide derivatization was prepared as follows: Add 200 µl of 2,3-napthalenedicarboxaldehyde (**122**, 10 mM solution in methanol), 200 µl of taurine (**123**, 50 mM solution in water), 60 µl of 7M ammonia in methanol, 200 µl of methanol and 140 µl of water. Take 80 µl of this stock solution in an Eppendorf tube and add 20 µl of cyanide (**82**) standard or enzymatic reaction sample. Incubate at room temp for 15-30 min and quench the reaction by adding 100 µl of water. Inject 10 µl of this derivatized sample on HPLC and/or LC-MS.

Phosphoenolpyruvate carboxylase (PEPC) mediated derivatization of CO_2:[94] All steps were performed in anaerobic chamber. Prepare all stock solutions in 100 mM KPi buffer, pH 7.5. Dissolve 1 mg of PEPC (activity 1.77 units/mg) in 200 µl of buffer, 10 mM phosphoenolpyruvate (PEP) solution in buffer, 10 mM $MgCl_2$ solution in buffer. To derivatize CO_2 generated during the NosL catalyzed reaction, include 15 µl of PEPC, 1 mM of PEP and 1 mM of $MgCl_2$ while setting up NosL assay. The derivatize the oxaloacetate for LC-MS analysis, filter the enzymatic reaction using 10 kda cut-off filters (VWR) and add 5 mM o-(2,3,4,5,6-pentafluorobenzyl)-hydroxylamine.HCl (PFBHA) and heat at 60 °C for 20 min.

HPLC parameters: An Agilent 1260 HPLC equipped with a quaternary pump was used. The system included a diode array UV-Vis detector and eluted compounds were detected by absorbance at 254 nm, 260 nm, 280 nm, 288 nm, 320 nm, 340 nm and 475 nm. The parameters for the fluorescence detector were: excitation at 270 nm and emission at 352 nm (to detect indole based products) or excitation at 418 nm and emission at 454 nm (to detect derivatized cyanide **124**). For analytical purposes, HPLC analysis was performed on a ZORBAX Eclipse XDB-C18 column (15 cm x 4.6 mm, 5 µm particles, Agilent Technologies, flow rate = 1 ml/min). Purification of the isotopes of L-tryptophan or its analogs, was carried out on semi-prep column from Supelcosil LC-18 (25 cm x 10 mm, 5 µm particles, Supelco, flow rate 2 ml/min). Typical injection volumes were in the range of 10-80µl. Data was processed using ChemStation ver. B.04.01 SP1 (Agilent technologies).

HPLC conditions:

A- Water

B- 100 mM Potassium phosphate buffer, pH 6.6 / 10 mM Ammonium acetate, pH 6.6

C- Methanol

HPLC method: (Flow rate 1 ml/min)

0 min – 100% B, 5 min – 100% B, 14 min – 7% A 70% B 23% C, 25 min – 25% A 0% B 75% C, 28 min - 25% A 0% B 75% C, 32 min - 100% B, 36 min - 100% B.

LC-MS parameters: LC-ESI-TOF-MS was performed using an Agilent 1260 HPLC system equipped with a binary pump and a 1200 series diode array detector followed by a MicroToF-Q II mass spectrometer (Bruker Daltonics) using an ESI source either in negative mode or positive mode. The samples were run either on LC-18-T column (15 cm x 3 mm, 3 µm particles, Supelco, flow rate 0.4 ml/min) or Poroshell 120 EC-C18 (10 cm x 3 mm, 2.7 µm particles, Agilent, flow rate 0.4 ml/min). Typical injection volumes were in the range of 10-80 µl. The data was processed using DataAnalysis 4.0 SP1 (Bruker Daltonics).

LC conditions: (Flow rate 0.4 ml/min)

A- 5 mM Ammonium acetate buffer, pH 6.6

B- 75% Methanol and 25% Water.

LC method: (for negative mode on MS)

0 min – 100% A, 7 min – 100% A, 10 min – 80% A 20% B, 27 min – 100% B, 29 min – 100% B, 30 min – 100% A, 40 min – 100% A.

LC method: (for positive mode on MS)

0 min – 100% A, 2 min – 100% A, 4 min – 80% A 20% B, 27 min – 100% B, 29 min – 100% B, 30 min – 100% A, 40 min – 100% A.

MS parameters: (for all LC-MS experiments)

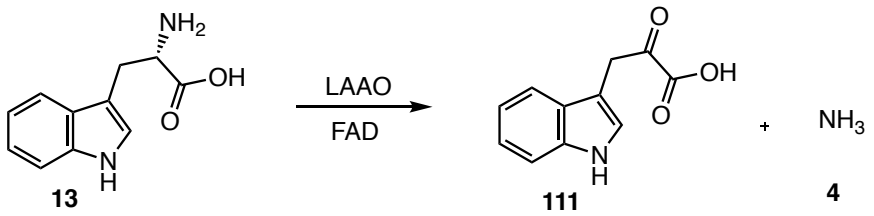
Capillary, -4500 V; capillary offset, -500 V; nebulizer gas, 3.0 bar; dry gas, 10 L/min; dry gas temperature, 200 °C; funnel 1 RF, 250.0 Vpp; funnel 2 RF, 300.0 Vpp; ISCID,

0.0 eV; hexapole RF, 200 Vpp; quadrapole ion energy, 3.0 eV; collision cell, collision energy, 8.0 eV; collision RF, 150.0 Vpp, transfer time, 80.0 µs; prepulse storage, 5.0 µs.

NMR analysis:

NMR spectra of all the synthetic samples and enzymatic reactions were recorded on a Bruker Avance III 400 MHz instrument.

Purity of substrate analogs: During our investigation, we found that NosL accepts D-stereoisomers of tryptophan and its analogs as substrates and generates different β-scission products. Thus, utmost care was taken to ensure L-stereoisomer of analogs is used wherever mentioned in this section. After synthesis, all analogs were purified on HPLC using a semi-prep column. The stereo-specificity of these analogs was tested by treating them with L-amino acid oxidase (LAAO) and D-amino acid oxidase (DAAO).[95] Complete consumption of each substrate analog when treated with L-amino acid oxidase confirmed that only L-stereoisomer of each compound was present (Figure 4.35). Additionally, no activity was detected with D-amino acid oxidase with any analog.[96]



13 **111** **4**

Figure 4.35: L-amino acid oxidase (LAAO) was used to test the stereospecific purity of substrate analogs. This enzyme specifically acts on the L-stereoisomer of amino acids.

Synthesis of L-tryptophan thiocarboxylate (115):[97]

Step 1: In a dry round bottom flask, dissolve 2 mmol of 2,5-dioxopyrrolidin-1-yl (((9H-fluoren-9-yl)methoxy)carbonyl)-L-tryptophanate (**143**, from Chem-Impex International Inc.) and 2 mmol of sodium hydrosulfide in dry ethanol and stir at room temperature for 8 h. Solvent was removed on rotovap, residue dissolved in 50 ml of water, acidified with 5% HCl up to pH = 3 and extracted twice with 30 ml of ethyl acetate. Combined ethyl acetate solutions were dried with anhydrous Na_2SO_4, separated and concentrated under reduced pressure to obtain crude product, which was additionally purified by column chromatography on silica gel with hexanes/ethyl acetate mixture 4:1-2:1, to give sticky oil 50-60% yield of **144**.

Step 2: Solution of 200 mg of (S)-2-((((9H-Fluoren-9-yl)methoxy)carbonyl)amino)-3-(1H-indol-3-yl)propanethioic S-acid (**144**) in 1 ml of 20% pyrrolidine in DMF was stirred at room temperature for 4 hr. The solvent was the removed under reduced pressure and the residue was subjected to flash chromatography purification using silica gel. Elution using ethyl acetate and methanol gave 50 mg of white solid (**115**, 50%). The compound was further purified using HPLC.

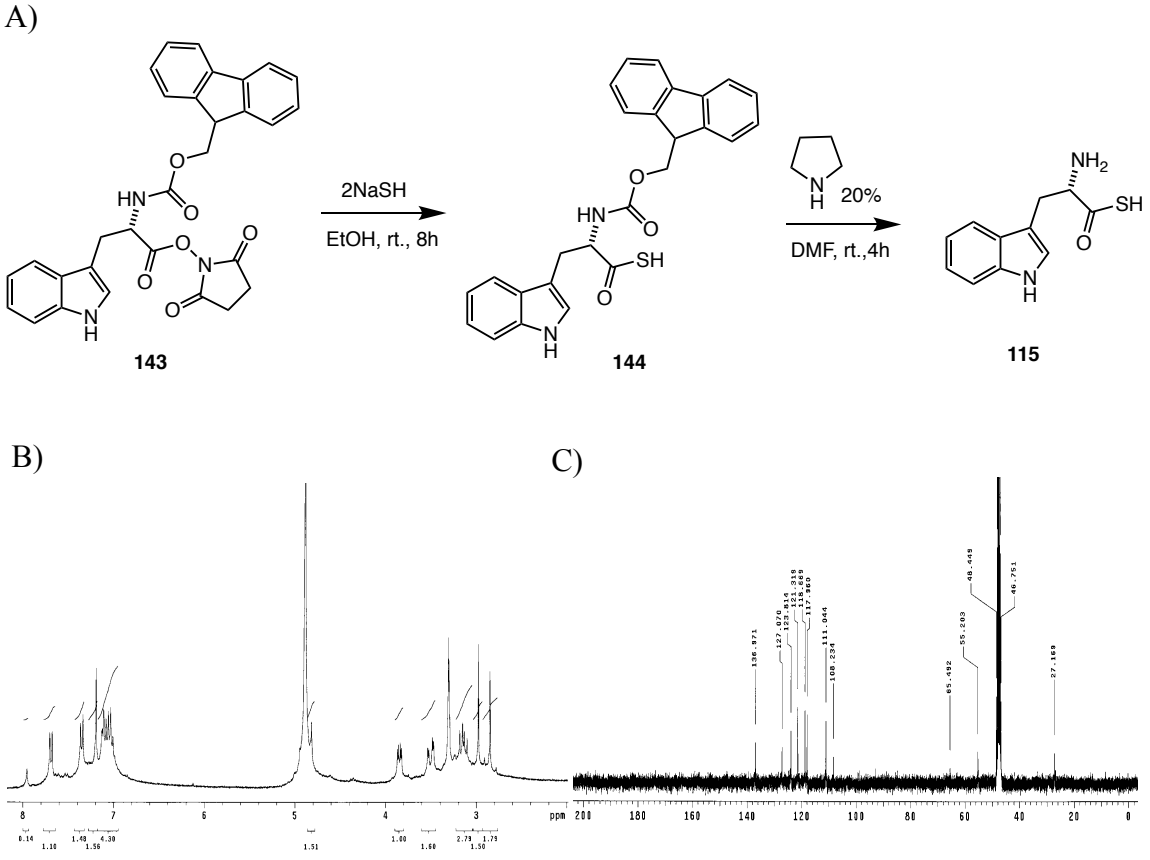


Figure 4.36: Synthesis of L-tryptophan thiocarboxylate (**115**). A) Synthetic scheme. B) ^{1}H-NMR of the product L-tryptophan thiocarboxylate (**115**). C) ^{13}C-NMR of the product **115**.

Synthesis of 2'-chloro-L-tryptophan (119):

2'-Chloro-L-tryptophan (**119**) was synthesized by a previously reported protocol with minor variations.[98] In a dry round bottom flask, 5.1 mmol of L-tryptophan methyl ester.HCl (**145**, 1.3 gm) was dissolved in 50 ml acetic acid and 15 ml formic acid. To this solution 5.2 mmol of N-chlorosuccinimide (**146**, 700 mg) was added. The reaction mixture was stirred at room temperature for 30 min and solvents were removed using

rotovap. The product **147** was purified by flash chromatography (silica, dichloromethane:methanol = 95:5). Isolated yield was ~50%. Since the C_2-chloro group is susceptible to acid/base hydrolysis, de-protection of the ester group was performed using α-chymotrypsin. 500 mg of **147** was dissolved in 10 ml of dioxane in a round bottom flask. A separate solution containing 50 mg of α-chymotrypsin (Type II, from bovine pancreas), 45 ml of water, 2 ml of 1M $NaHCO_3$ and 10 ml dioxane, was added to this solution in a round bottom flask. The reaction mixture was stirred for ~6 hr at room temperature which was further frozen using Liq. N_2 and lyophilized overnight. The residual mixture was suspended in methanol and filtered through celite. The residue was washed with methanol and the solvent was removed using rotovap to obtain **119**. Isolated yield was ~90%. NMR was recorded in CD_3OD. Compound **119** was further purified using HPLC.

For synthesis of 2'-bromo-L-tryptophan (**121**), procedure described above was adopted with L-tryptophan methyl ester.HBr and N-bromosuccinimide.

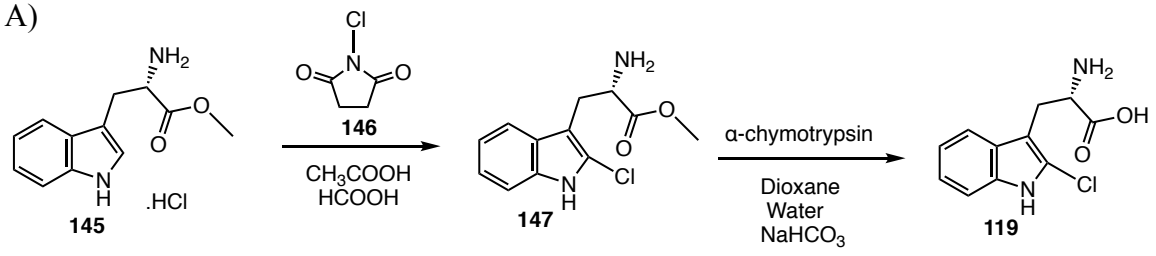


Figure 4.37: Synthesis of 2'-chloro-L-tryptophan (**119**). A) Synthetic route. B) ^{1}H-NMR of the product 2'-chloro-L-tryptophan (**119**). C) ^{13}C-NMR of the product **119**.

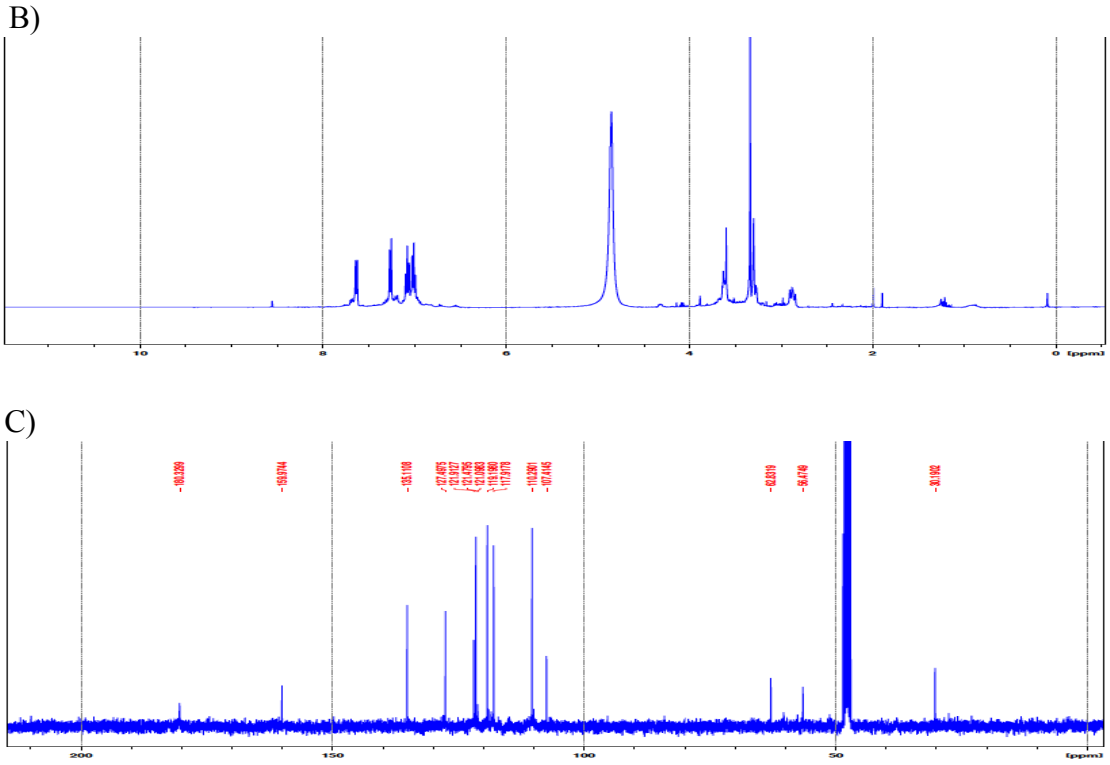


Figure 4.37 continued.

Synthesis of 2'-methyl-L-tryptophan (130)

2'-methyl-D,L-tryptophan (**130a**) was obtained commercially from Chem-Impex international Inc. This compound was acetylated using acetic anhydride by heating at 60 °C for 30 min in acetic acid (Figure 4.38). The solvent was removed using rotovap and solid was dissolved in KPi buffer, pH 8. Amano acylase was added to this solution and it was stirred for 2 hr at 50 °C. Stereospecific 2'-methyl-L-tryptophan (**130**) was obtained by HPLC purification.

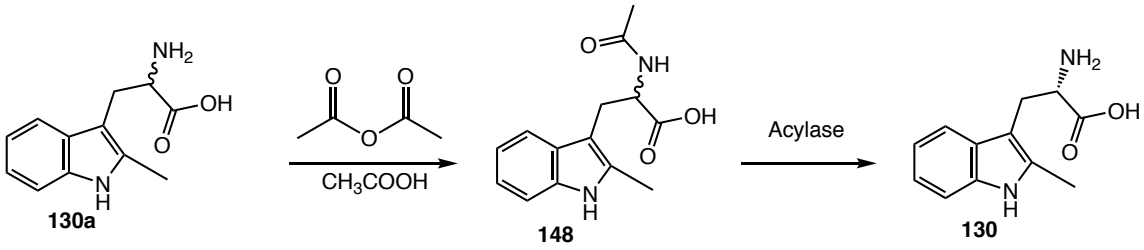


Figure 4.38: Scheme depicting synthesis of 2'-methyl-L-tryptophan (**130**) from 2'-methyl-D,L- tryptophan (**130a**).

5. TRYPTOPHAN LYASE (NOSL): A CORNUCOPIA OF 5′-DEOXYADENOSYL RADICAL MEDIATED TRANSFORMATIONS*

5.1 The reactive 5′-deoxyadenosyl radical (6)

In the biological world, complex transformations often involve radical intermediates. Enzymes from different families like the adenosyl cobalamin dependent enzymes,[99] cytochrome P450 dependent enzymes,[66] non-heme iron dependent enzymes[100] and radical S-adenosyl-L-methionine (SAM)[12] dependent enzymes catalyze these radical mediated reactions. Amongst these enzyme families, adenosyl cobalamin dependent enzymes and radical SAM enzymes utilize the 5′-deoxyadenosyl radical (**6**) as a potent oxidant. The bond dissociation energy of the C_5'-H in 5′-deoxyadenosine is ~433 kJ/mol, making it the most reactive organic radical in living organisms.[101] To avoid inactivating the enzyme with this highly reactive radical, its reactivity must be tightly controlled at the active site. It is generally assumed that one of the control mechanisms involves formation of the 5′-deoxyadenosyl radical in close proximity to the abstracted hydrogen atom.[102] In this section we describe studies with enzyme and substrate analogs that reveal unanticipated mobility of the radicals formed at the active site of NosL.

---

*Reprinted in parts with permission from "Tryptophan Lyase (NosL): Mechanistic Insights from Substrate Analogues and Mutagenesis" by Bhandari, D. M., Xu, H., Nicolet, Y., Fontecilla-Camps, J. C., & Begley, T. P. *Biochemistry*, **2015**, *54*(31), 4767-4769. Copyright 2015 American Chemical Society and from "Tryptophan Lyase (NosL): A Cornucopia of 5′-Deoxyadenosyl Radical Mediated Transformations" by Bhandari, D. M., Fedoseyenko, D., & Begley, T. P. *J. Am. Chem. Soc.* **2016**, *138* (50), 16184-16187. Copyright 2016 American Chemical Society.

5.2 Results and discussion

5.2.1 Reaction of NosL with N-methyl-L-tryptophan (149)

NosL catalyzes the demethylation of N_{α}-methyl-L-tryptophan (**149**), followed by the conversion of L-tryptophan (**13**) to 3-methyl-2-indolic acid (**15**) and 3-methylindole (**92**) (Figure 5.1).

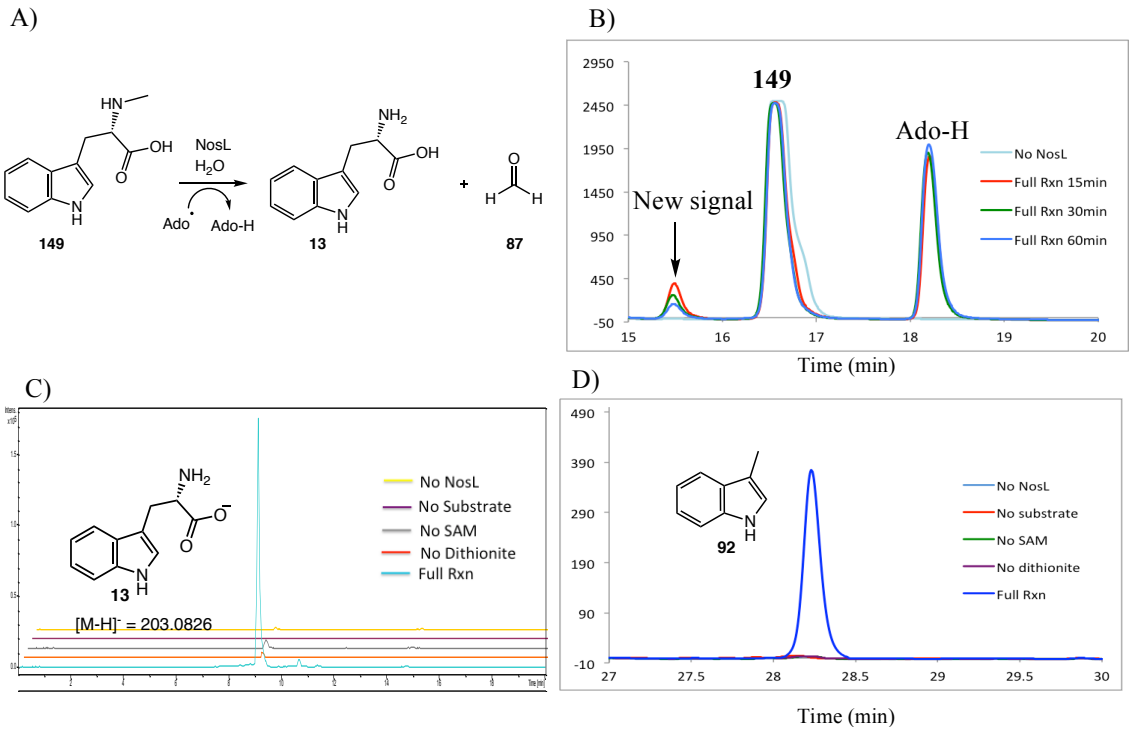


Figure 5.1: Analysis of the NosL reaction with N_{α}-methyl-L-tryptophan (**149**). A) Scheme showing demethylation of **149** catalyzed by NosL. B) HPLC analysis of the time-course showed formation of a new signal at 15.5 min, which was consumed as the duration of the enzymatic reaction was increased. C) LC-MS analysis showing EIC [M-H]$^-$ = 203.0 Da confirming the identity of the new signal as L-tryptophan (**13**) (only observed in the 'full reaction'). D) HPLC analysis showing formation of 3-methylindole (**92**) observed only in the 'full reaction' sample. E) LC-MS analysis of the enzymatic reaction performed with the FldA/R system showing EIC of [M-H]$^-$ = 174.0 Da confirming formation of 3-methyl-2-indolic acid (**15**) in the 'full rxn' sample. F) MS of **15** generated during enzymatic reaction.

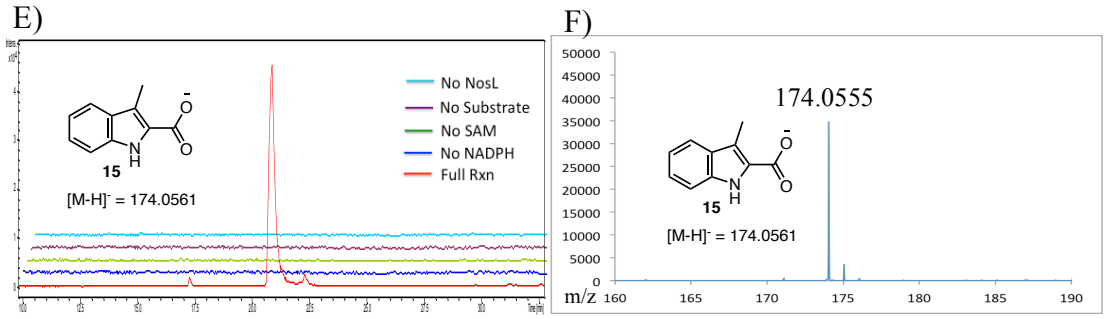


Figure 5.1 continued.

To study the mechanism of demethylation, the enzymatic reaction was performed in 95% D_2O buffer. However, low deuterium incorporation was observed in 5′-deoxyadenosine suggesting that the H-atom abstraction position is not the amino group. When the reaction was carried out using N_α-methyl-d_3-L-tryptophan (**149a**), deuterium transfer from the methyl group to 5′-deoxyadenosine (**93a**) was observed (Figure 5.2). This demonstrates that the 5′-deoxyadenosyl radical (**6**) abstracts a H-atom from the methyl group instead of the amino group as previously seen in the native reaction.

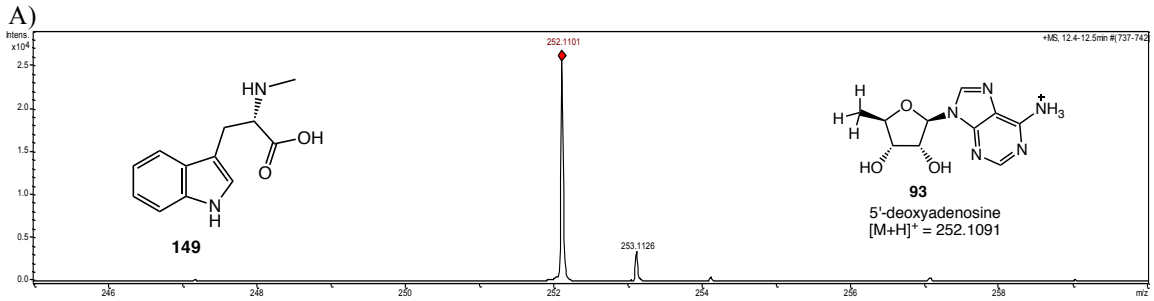


Figure 5.2: MS of 5′-deoxyadenosine (**93**) generated. A) 5′-Deoxyadenosine (**93**) generated in the ‘full reaction’ with N_α-methyl-L-tryptophan (**149**). B) 5′-Deoxyadenosine (**93a**) generated in the ‘full reaction’ with N_α-methyl-d_3-L-tryptophan (**149a**).

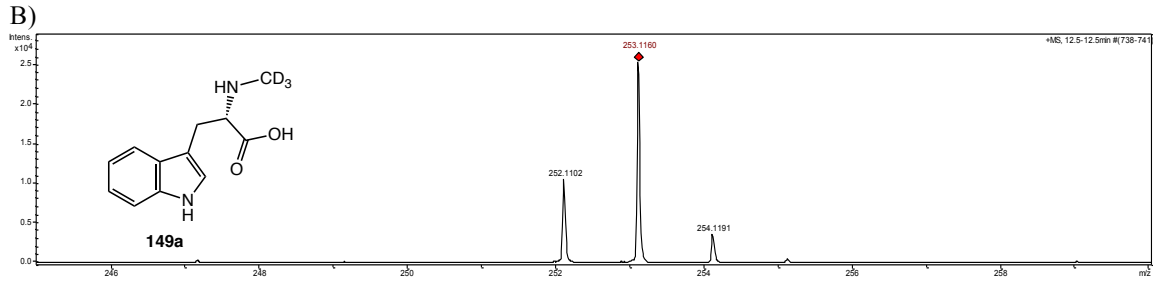


Figure 5.2 continued.

A mechanistic proposal for the NosL-dependent demethylation reaction of **149** is shown in Figure 5.3. The 5′-deoxyadenosyl radical (**6**) abstracts a H-atom from the methyl group of **149** generating **150**. Intermediate **150** loses an electron to the [4Fe-4S] generating **151** which is be rapidly hydrolyzed to L-tryptophan (**13**).

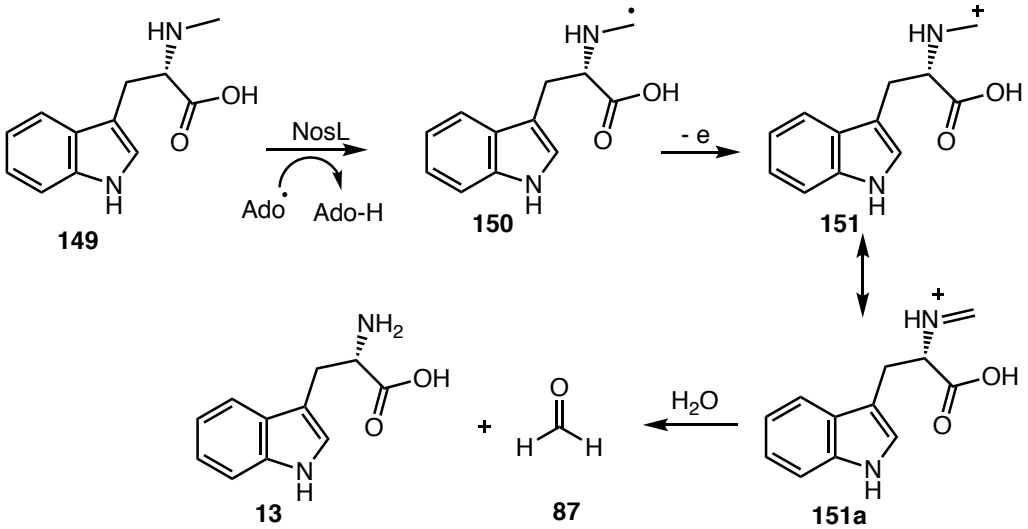


Figure 5.3: Mechanistic proposal for the NosL catalyzed demethylation of N_{α}-methyl-L-tryptophan (**149**).

5.2.2 Reaction of NosL with N-cyclopropyltryptophan (152)

N_α-cyclopropyltryptophan (**152**) was also investigated as a probe for the nitrogen centered radical, but this molecule was not a substrate for the WT enzyme. Reasoning that this might be due to the size of the cyclopropyl ring, the active site was enlarged by construction of the Y90A variant with the goal of using this analog to determine the relative rate of the β-bond scission of the amino radical. HPLC analysis of the NosL Y90A-catalyzed reaction with **152** revealed the formation of indole-3-pyruvic acid (**111**, Figure 5.4).

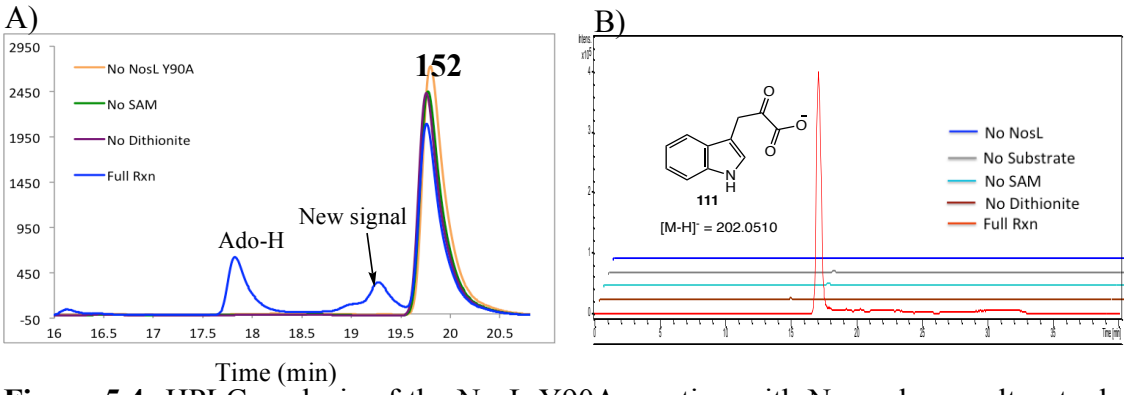


Figure 5.4: HPLC analysis of the NosL Y90A reaction with N_α-cyclopropyltryptophan (**152**). A) HPLC analysis at 280 nm indicates formation of a new signal in the 'Full Rxn' sample. B) LC-MS analysis showing EIC $[M-H]^-$ = 202.0 Da indicating the formation of indole-3-pyruvic acid (**111**) in the 'Full Rxn' sample. C) MS of indole-3-pyruvic acid (**111**) generated during the enzymatic reaction. D) LC-MS analysis showing EIC $[M-H]^-$ = 202.0 Da and co-injection with the authentic standard of **111** confirming its formation. E) Reaction scheme showing the derivatization of cyclopropylamine (**153**) with NBD-F (**154**). F) LC-MS analysis after NBD-F treatment showing EIC $[M-H]^-$ = 219.0 Da and 221 Da suggesting the formation of the adduct **155** in the 'Full Rxn' sample. G) MS of **155** generated after NBD-F derivatization of the enzymatic reaction. H) LC-MS analysis after NBD-F treatment showing EIC $[M-H]^-$ = 219.0 Da and co-injection with NBD-F derivatized authentic standard of **153** confirming formation of **155**.

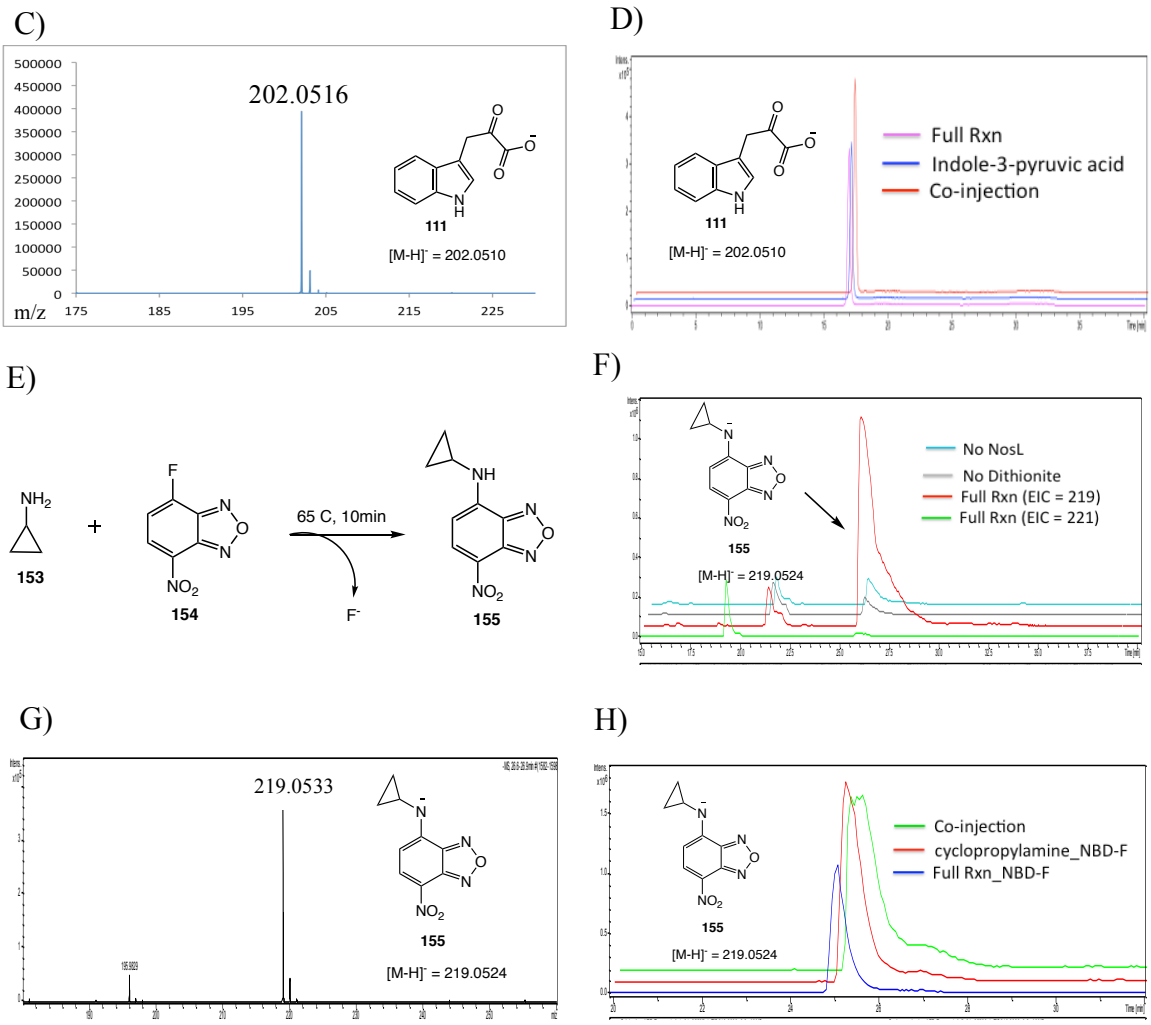


Figure 5.4 continued**.**

LC-MS analysis of the reaction mixture after treatment with 4-fluoro-7-nitrobenzofurazan (NBD-F, **154**)[103] demonstrated formation of the cyclopropylamine-NBD adduct (**155**) with no evidence for the anticipated cyclopropyl ring opening.

Additionally, deuterium incorporation into 5′-deoxyadenosine (**93**) was not observed when the reaction was carried out in 95% D_2O buffer. These results suggested that the amino group is not the site of the initial H-atom abstraction. To identify the

position of H-atom abstraction from the substrate, we synthesized **152a** deuterated at the C_α position. When this C_α-deuterated **152a** isotopologue was used for the enzymatic reaction with NosL Y90A, analysis of the 5′-deoxyadenosine (**93**) generated during the enzymatic reaction showed incorporation of a single deuterium (Figure 5.5).

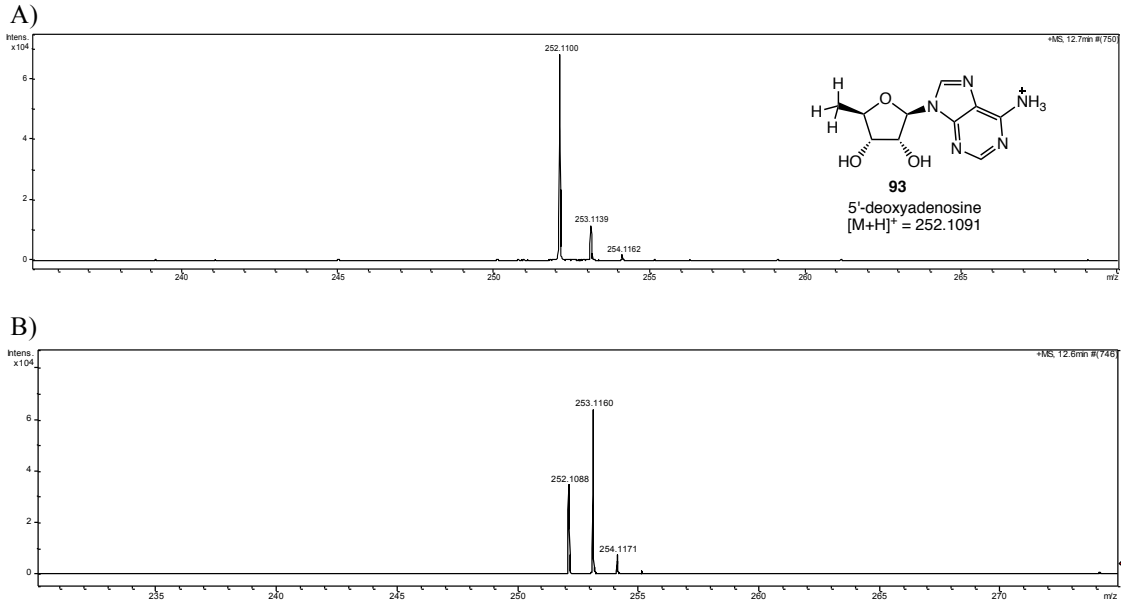


Figure 5.5: Analysis of the 5'-deoxyadenosine (**93**) generated during the NosL reaction with **152**. A) MS of **93** generated during the reaction of NosL Y90A with **152** in 95% D_2O buffer. B) MS of **93** generated during the reaction of NosL Y90A with 1**52a**.

A mechanistic proposal consistent with these observations is shown in Figure 5.6. H-atom abstraction from the C_α position of **152** gives **156**. Amine deprotonation followed by electron transfer back to the cluster gives **158**, which undergoes hydrolysis to give the observed products **111** and **153**.

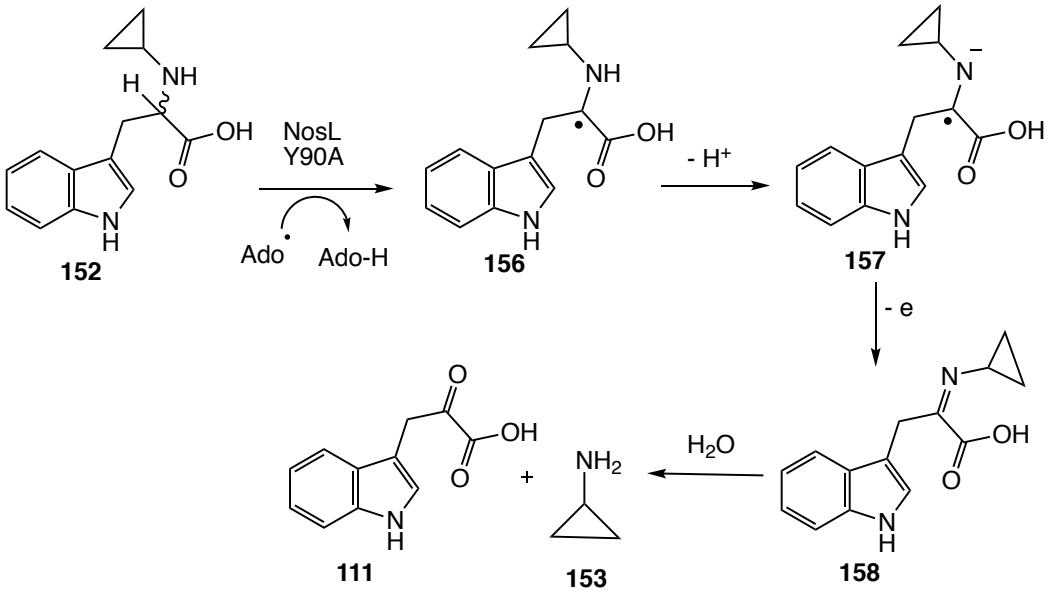


Figure 5.6: Mechanistic proposal for the NosL Y90A-catalyzed reaction of N_{α}-cyclopropyltryptophan (**152**).

This experiment completes the series shown in Figure 5.7 in which the active site 5′-deoxyadenosyl radical (**6**) of NosL abstracts a hydrogen atom not only from the amine of the substrate (**13**), but also from the methyl group of **149** and Cα of **152**.

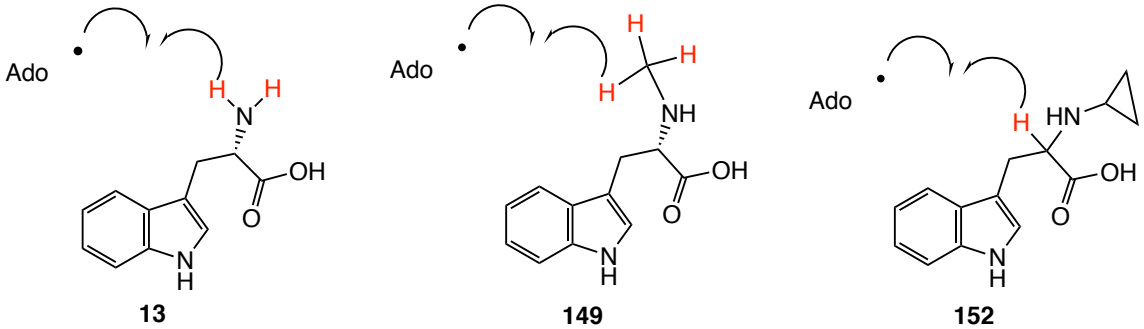


Figure 5.7: The 5′-deoxyadenosyl radical (**6**) in the active site of NosL shows relaxed regiospecificity with substrate analogs **149** and **152**.

5.2.3 Reaction of NosL with D,L-indole-3-lactic acid (159)

NosL is the first example in radical SAM enzymology where a H-atom abstraction from the amino-group has been unequivocally demonstrated. The N-H bond dissociate energy (BDE) is ~425 kJ/mol while that of C_5'-H in 5′-deoxyadenosine is ~433 kJ/mol.[101] We were interested in investigating if the 5′-deoxyadenosyl radical (**6**) is capable of abstracting a H-atom from the alcoholic O-H which has BDE of ~440 kJ/mol (phenolic O-H has BDE of ~360 kJ/mol making this H-atom abstraction highly favorable). To experimentally test this possibility, the NosL reaction was performed using D,L-indole-3-lactic acid (**159**) as substrate. As shown in Figure 5.8, **159** acts as a substrate for NosL and generates **111** and **108**. This demonstrates that the 5′-deoxyadenosyl radical (**6**) can abstract a H-atom from the alcoholic O-H group.

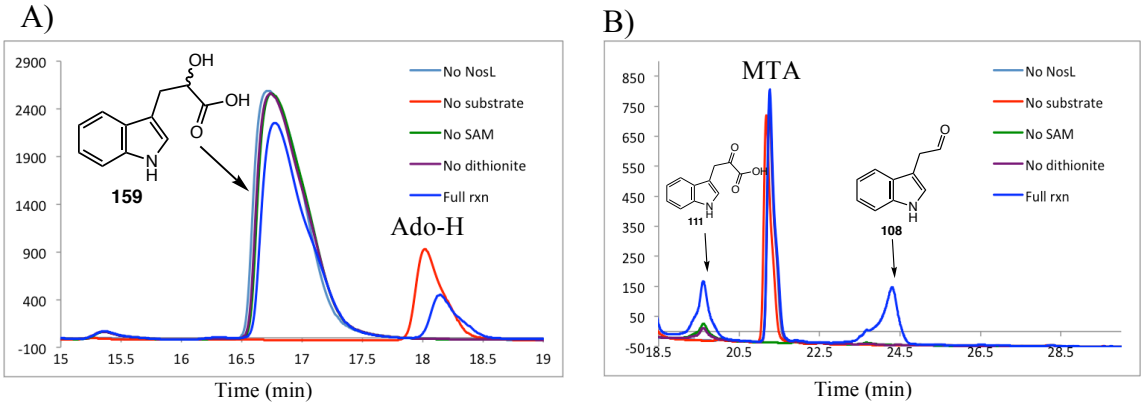


Figure 5.8: Analysis of the NosL enzymatic reaction with D,L-indole-3-lactic acid (**159**). A) HPLC analysis at 280 nm showed consumption of substrate only in the 'full reaction' sample. B) HPLC analysis at 280 nm showing two new signals corresponding to **111** and **108**. C) LC-MS analysis showing EIC $[M-H]^-$ = 202.0 Da confirming formation of **111** in the 'full reaction' sample. D) MS of **111** generated. E) LC-MS analysis showing EIC $[M-H]^-$ = 158 Da indicating formation of **108** in the 'full reaction' sample. F) MS of **108** generated during the NosL-catalyzed reaction.

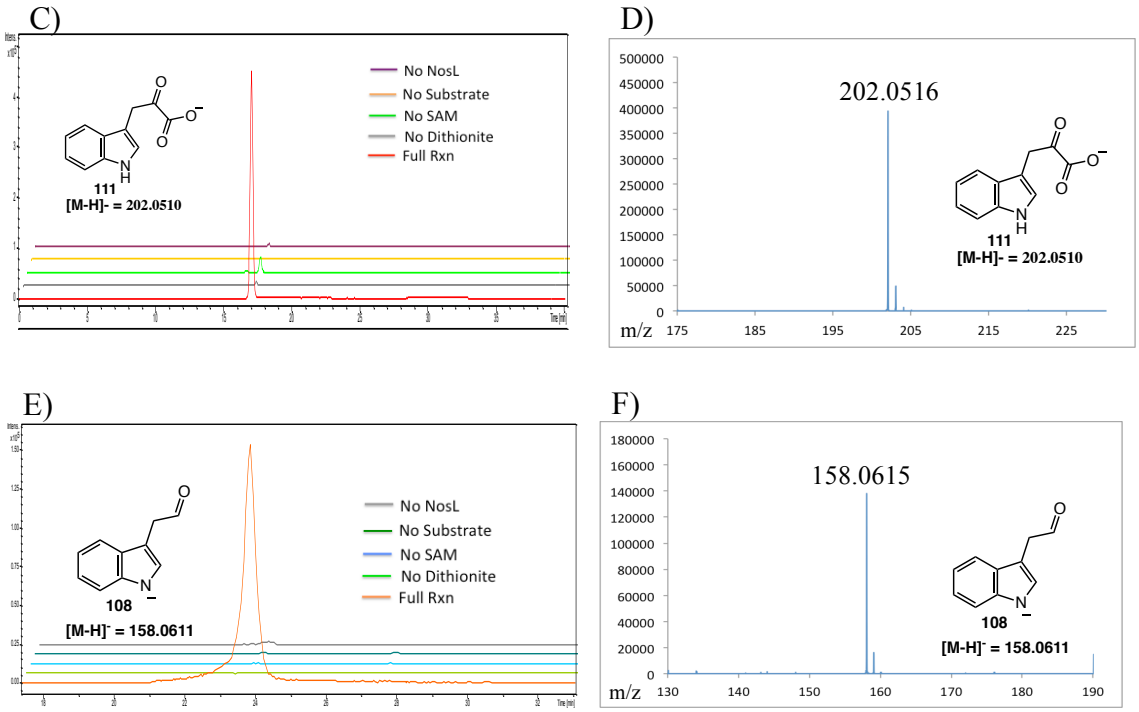


Figure 5.8 continued.

To further confirm formation of indole-3-acetaldehyde (**108**), the 'full reaction' sample was reduced with $NaBH_4$ generating tryptophol (**113**) (Figure 5.9). This sample was then co-injected with the authentic standard of **113** demonstrating that **108** is generated when the NosL reaction is performed in the presence of D,L-indole-3-lactic acid (**159**).

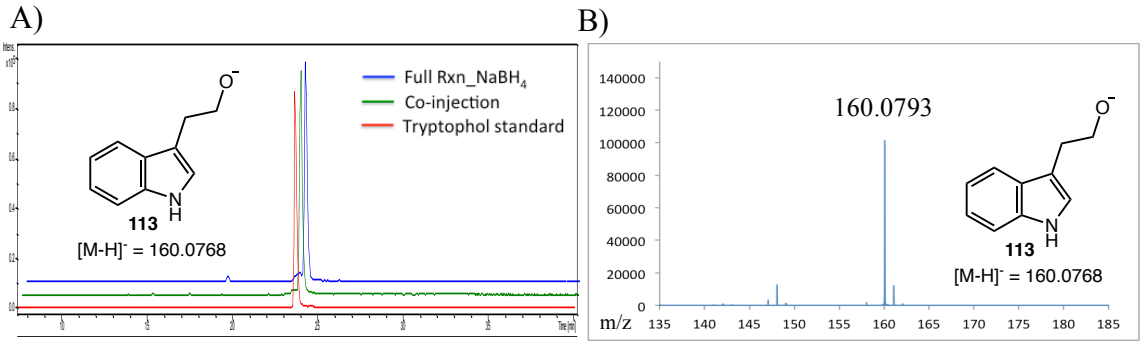


Figure 5.9: Analysis of tryptophol (**113**) generated by $NaBH_4$ reduction of indole-3-acetaldehyde (**108**) which is produced during the NosL reaction with D,L-indole-3-lactic acid (**159**). A) LC-MS analysis showing EIC $[M-H]^-$ = 160.0 Da for the $NaBH_4$ reduced sample of the 'full reaction'. The sample was co-injected with authentic standard of **113** confirming the identity of the new signal. B) MS of **113** generated after treatment of the 'Full reaction' sample with $NaBH_4$.

A mechanistic proposal for the NosL catalyzed conversion of D,L-indole-3-lactic acid (**159**) to indole-3-pyruvic acid (**111**) and indole-3-acetaldehyde (**108**) is outlined in Figure 5.10. The 5′-deoxyadnosyl radical (**6**) abstracts a H-atom from the alcoholic O-H of **159** generating **160**. Intermediate **160** can undergo decarboxylation followed by the loss of electron to generate **108** or it can be deprotonated at Cα and the subsequent loss of electron will form **111**.

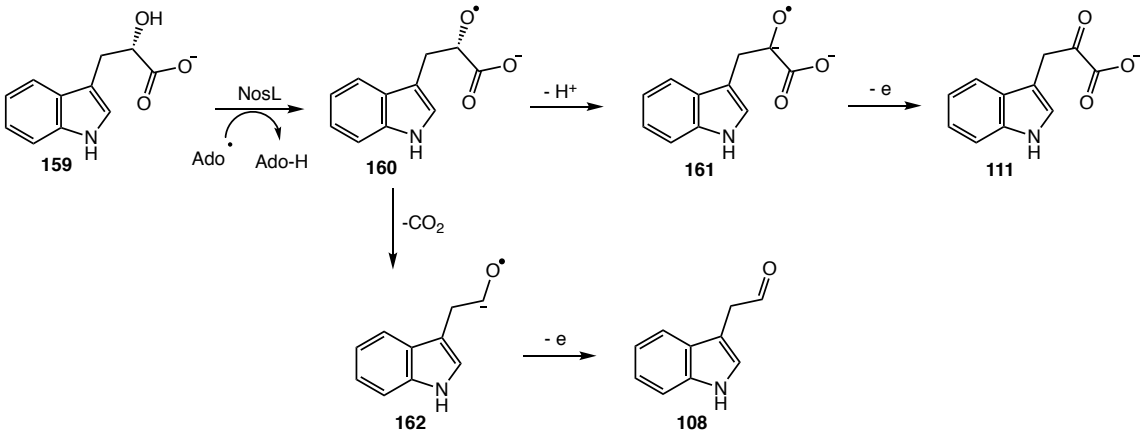


Figure 5.10: Mechanistic proposal for the NosL-catalyzed conversion of **159** into indole-3-pyruvic acid (**111**) and indole-3-acetaldehyde (**108**).

5.2.4 Reaction of NosL with methylene analog of tryptophan[104-105]

The MqnE-catalyzed formation of aminofutalosine (**164**) in menaquinone biosynthesis is a recently discovered exception to the general rule that H-atom abstraction is the first step in all reactions catalyzed by radical SAM enzymes (Figure 5.11).[106] In this system, the substrate radical is formed by addition of the 5′-deoxyadenosyl radical (**6**) to the double bond of **163** to give a radical intermediate that then undergoes rearrangement to give **164**.

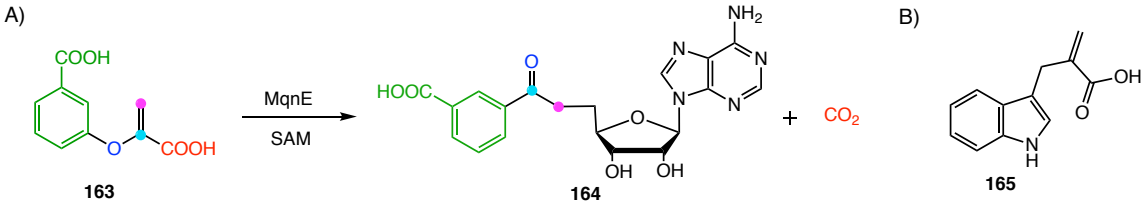


Figure 5.11: Addition of the 5′-deoxyadenosyl radical to the alkene group. A) MqnE-catalyzed addition of the 5′-deoxyadenosyl radical to **163** leading to the formation of aminofutalosine (**164**). B) Methylene analog of tryptophan (**165**).

The regiochemical tolerance of the 5′-deoxyadenosyl radical (**6**) at the active site of NosL suggested that NosL might catalyze a radical addition to the methylene analog of tryptophan (**165**).[107] When **165** was treated with NosL, three new products were observed by LC-MS analysis (Figure 5.12B). To reduce the unfavorable interaction between the methylene group in **165** and Arg323, we next investigated the reaction of **165** with NosL R323K. This variant showed ~10-fold enhancement in the formation of all three adducts (Figure 5.12D). It was not possible to characterize the reaction products by NMR because of their instability during purification.

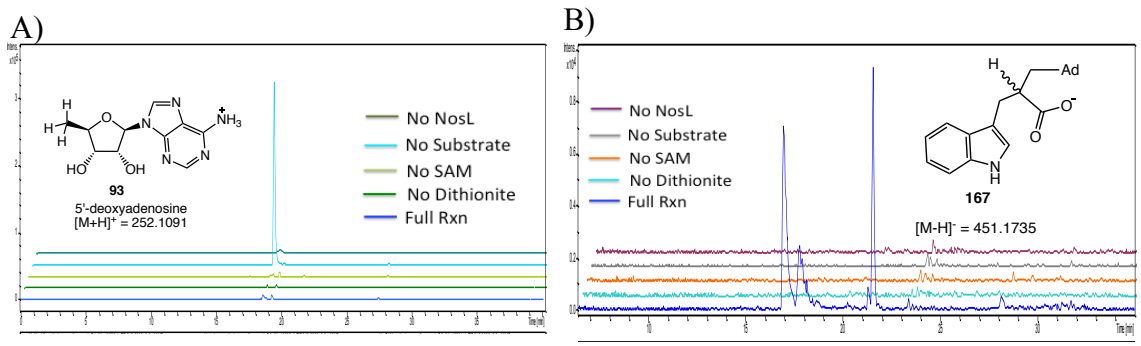


Figure 5.12: Analysis of the NosL reaction with the methylene analog **165**. A) LC-MS analysis showing EIC $[M-H]^-$ = 250.0 Da confirming the absence of **93** in the 'Full Rxn' sample with NosL WT. B) LC-MS analysis showing EIC $[M-H]^-$ = 451.1 Da suggesting the formation of three isomeric products corresponding to the mass of the adduct resulting from addition of the 5′-deoxyadenosyl radical to the **165** (one isomer shown). C) MS of the Peak 1 product generated in the 'Full Rxn' sample. D) HPLC analysis at 280 nm showing the enhanced formation of the three products when **165** is reacted with NosL R323K. E) LC-MS analysis showing EIC $[M-H]^-$ = 451.1 Da confirming that these three signals correspond to the mass of the adducts seen with NosL WT. F) MS of the Peak 1 product **167** generated in 'Full Rxn' sample with NosL R323K.

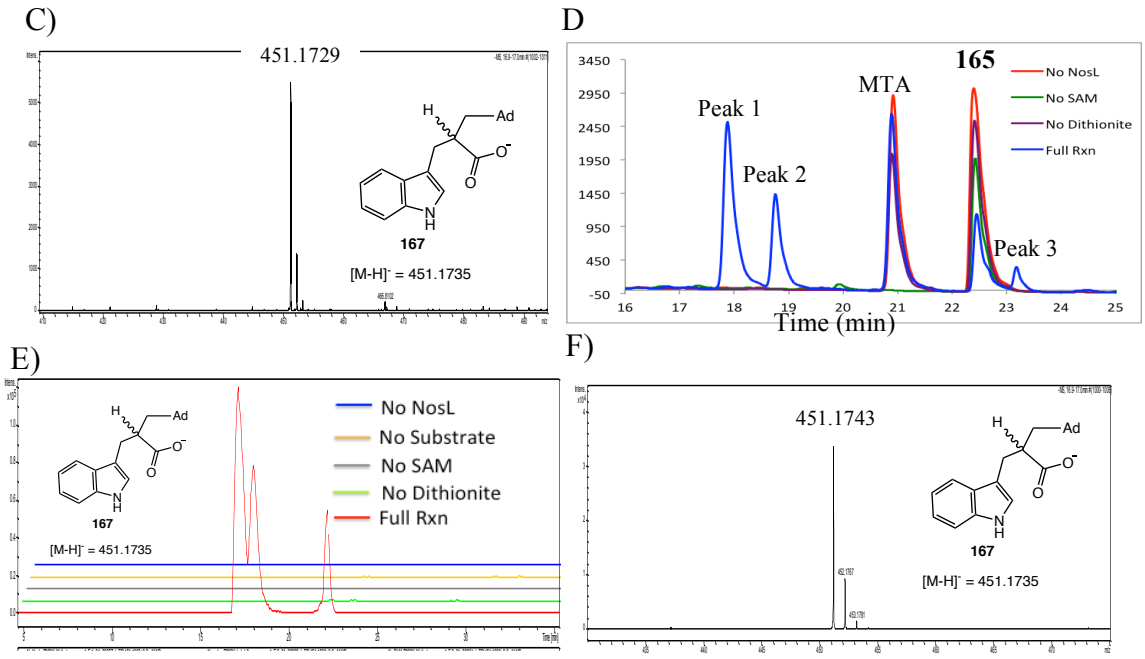


Figure 5.12 continued.

The three products with the identical m/z values are consistent with the expected mass of adducts between the substrate **165** and the 5′-deoxyadenosyl radical (**6**) suggesting that the latter radical added to both carbons of the double bond of **165** (Figure 5.13). This addition will be facially specific and determined by the active site structure. Quenching of **166** uses two different hydrogen atom donors and gives **167a, 167b** as a mixture of two epimers (path a). Quenching of radical **168** will give **169** as a single isomer (path b).

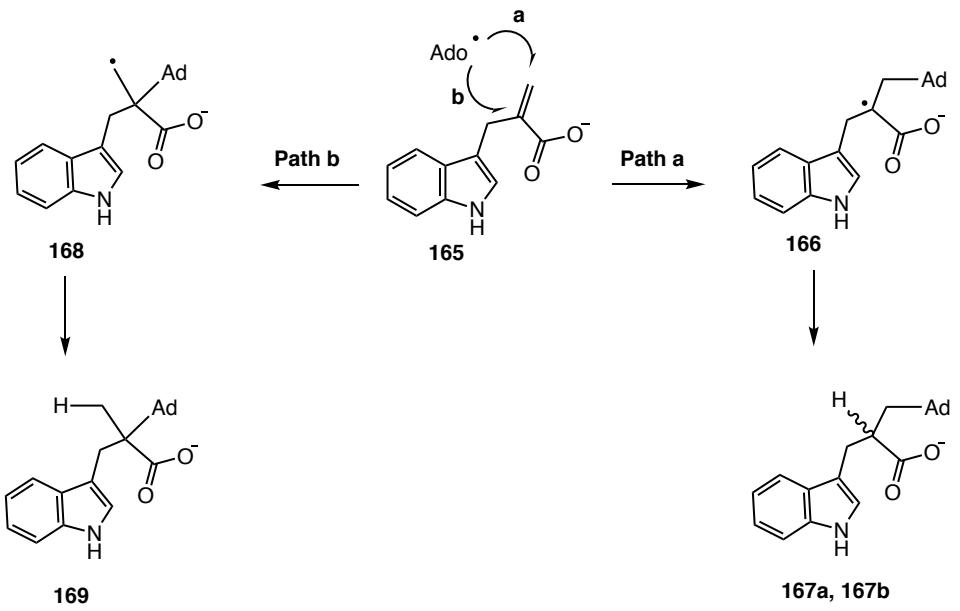


Figure 5.13: Addition of the 5′-deoxyadenosyl radical (**6**) to the double bond of indole methyl acrylate **165** during the NosL catalyzed reaction.

5.2.5 Reaction of NosL with indole-3-pyruvic acid (111)

In search of a tryptophan analog that would give more stable products, indole-3-pyruvic acid (**111**) was next tested using NosL R323K. HPLC analysis of the resulting reaction mixture indicated the absence of 5′-deoxyadenosine (**93**) and the formation of a new compound eluting at 24 min (Figure 5.14). LC-MS analysis of this product gave a mass that was consistent with **22**, generated by 5′-deoxyadenosyl radical addition to the carbonyl followed by radical induced decarboxylation of **20** and electron transfer back to the cluster (Figure 5.14A).

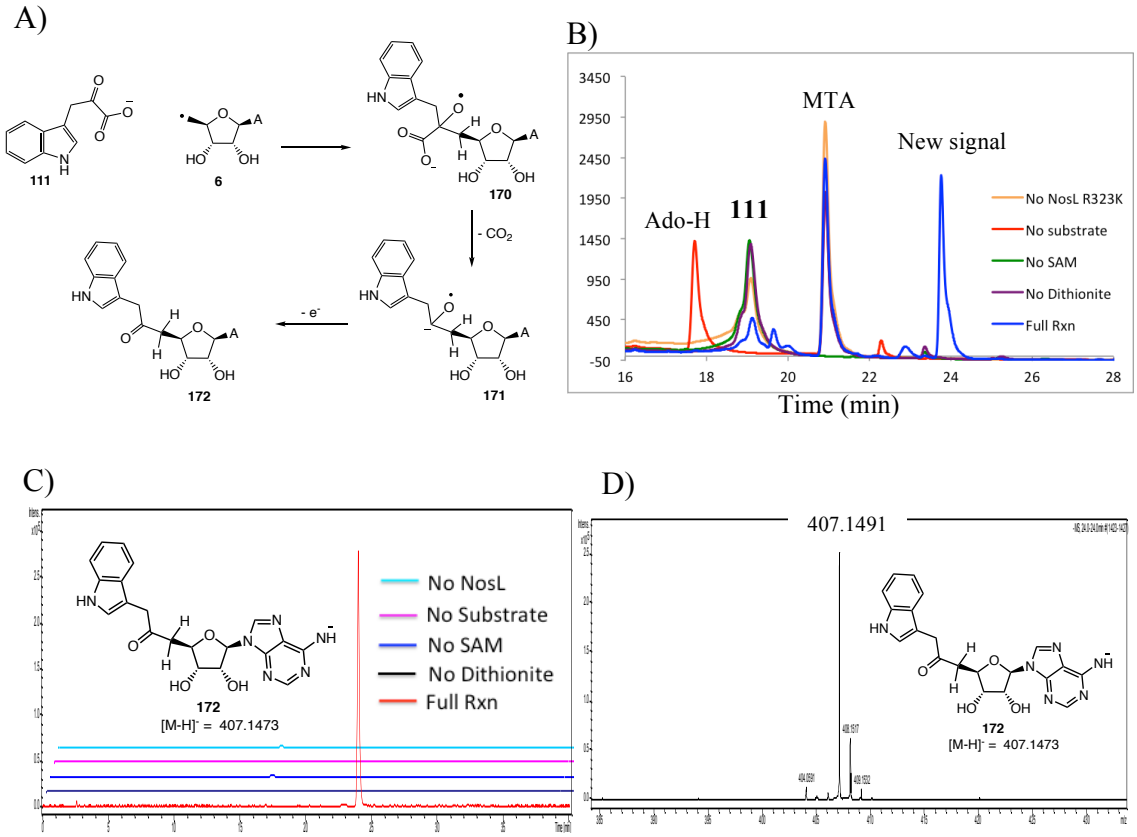


Figure 5.14: Reaction of NosL R323K with indole-3-pyruvic acid (**111**). A) Mechanistic proposal for the formation of **172** by addition of the 5′-deoxyadenosyl radical (**6**) to indole-3-pyruvic acid (**111**). B) HPLC analysis at 280 nm of the NosL R323K reaction with **111** showing consumption of substrate and formation of a new product only in the 'Full Rxn' sample. C) LC-MS analysis showing EIC $[M-H]^-$ = 407.1 Da suggesting the formation of an adduct **172**. C) MS of the reaction product **172** in negative mode.

The product **172** was stable and was purified by HPLC using 10 mM ammonium acetate as the elution buffer. The collected sample was lyophilized and dissolved in d_6-DMSO for NMR analysis (Figure 5.15).

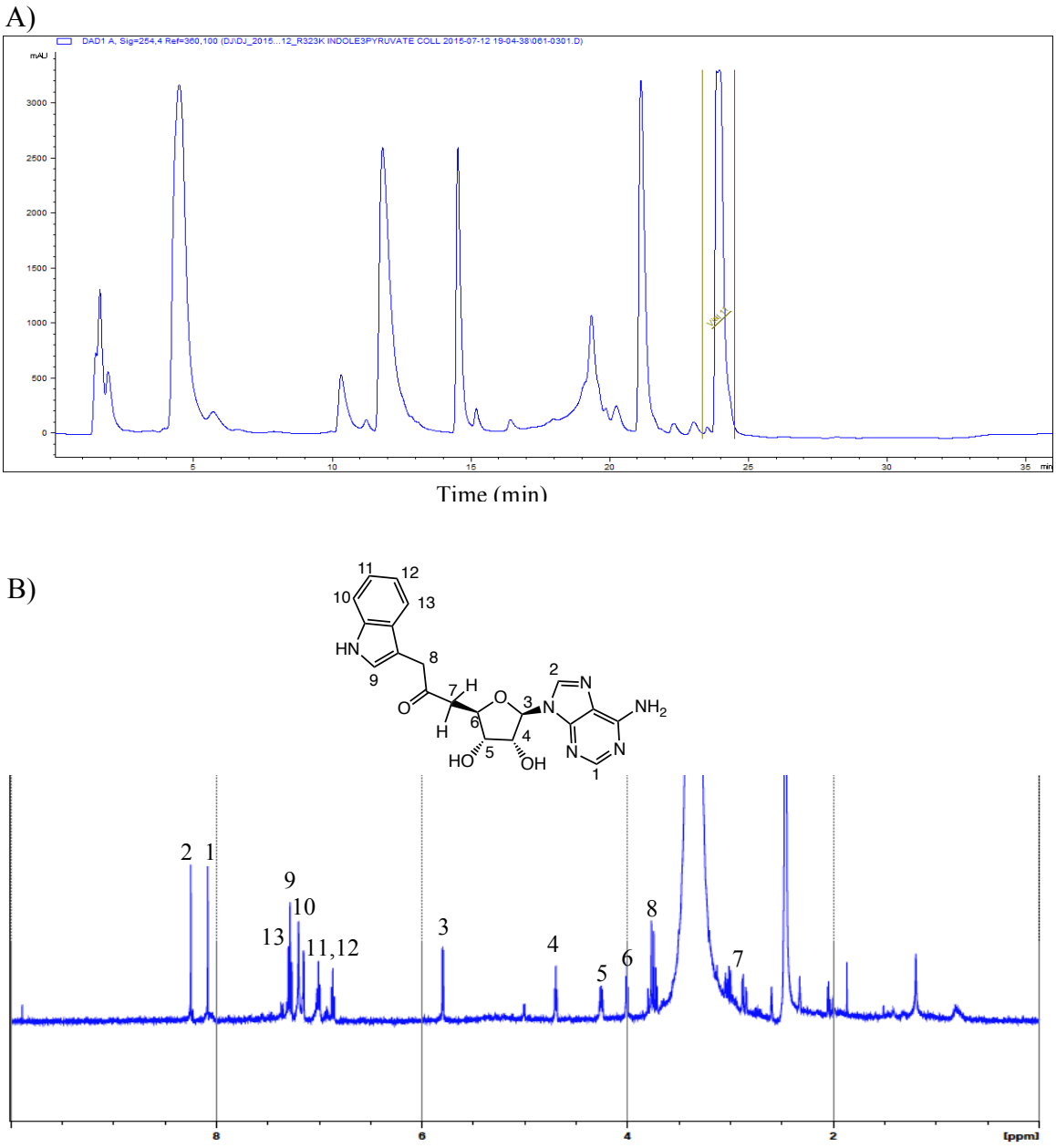
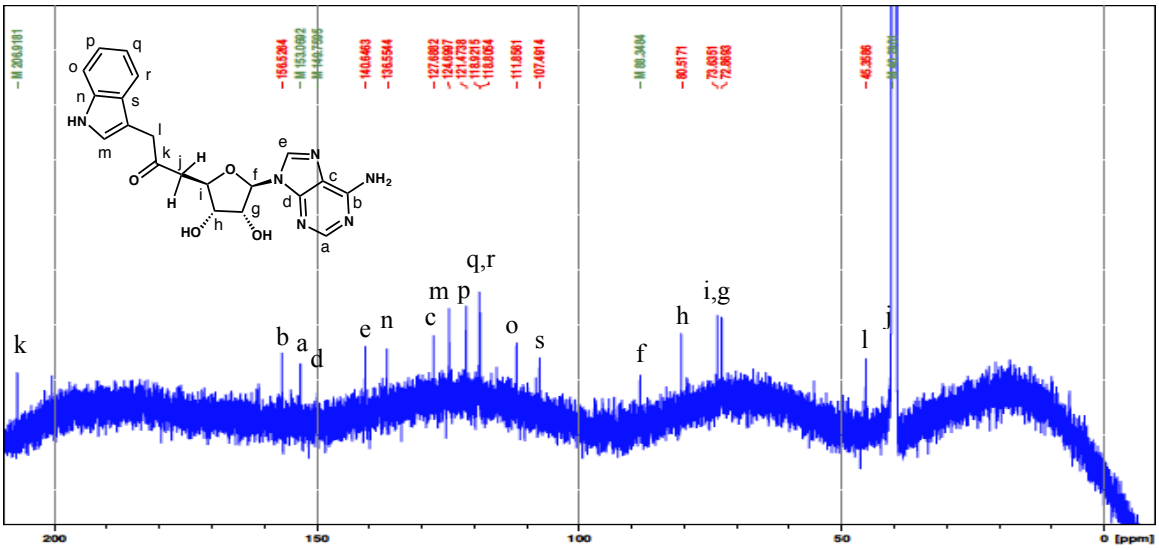


Figure 5.15: Characterization of the product generated by the reaction of NosL R323K with indole-3-pyruvic acid (**111**). A) HPLC analysis at 260 nm showing the collection of the signal corresponding to the product **172**. B) ^{1}H-NMR of the collected product **172** in d_6-DMSO. C) ^{13}C-NMR of compound **172**. D) ^{1}H-^{13}C HSQC-NMR of product **172**. E) ^{1}H-^{1}H COSY-NMR of product **172.** F) DEPT-135 NMR spectrum of compound **172**. CH_2 are shown negative while CH and CH_3 are positive.

C)



D)

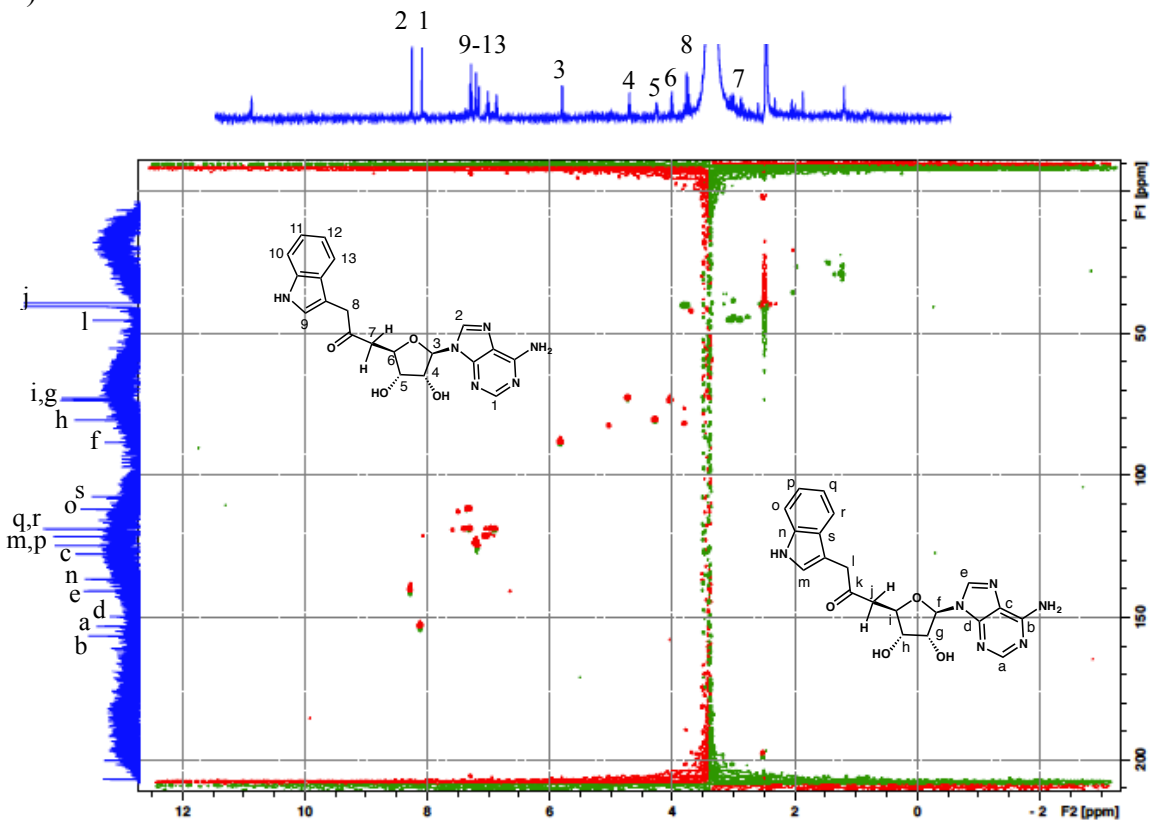


Figure 5.15 continued.

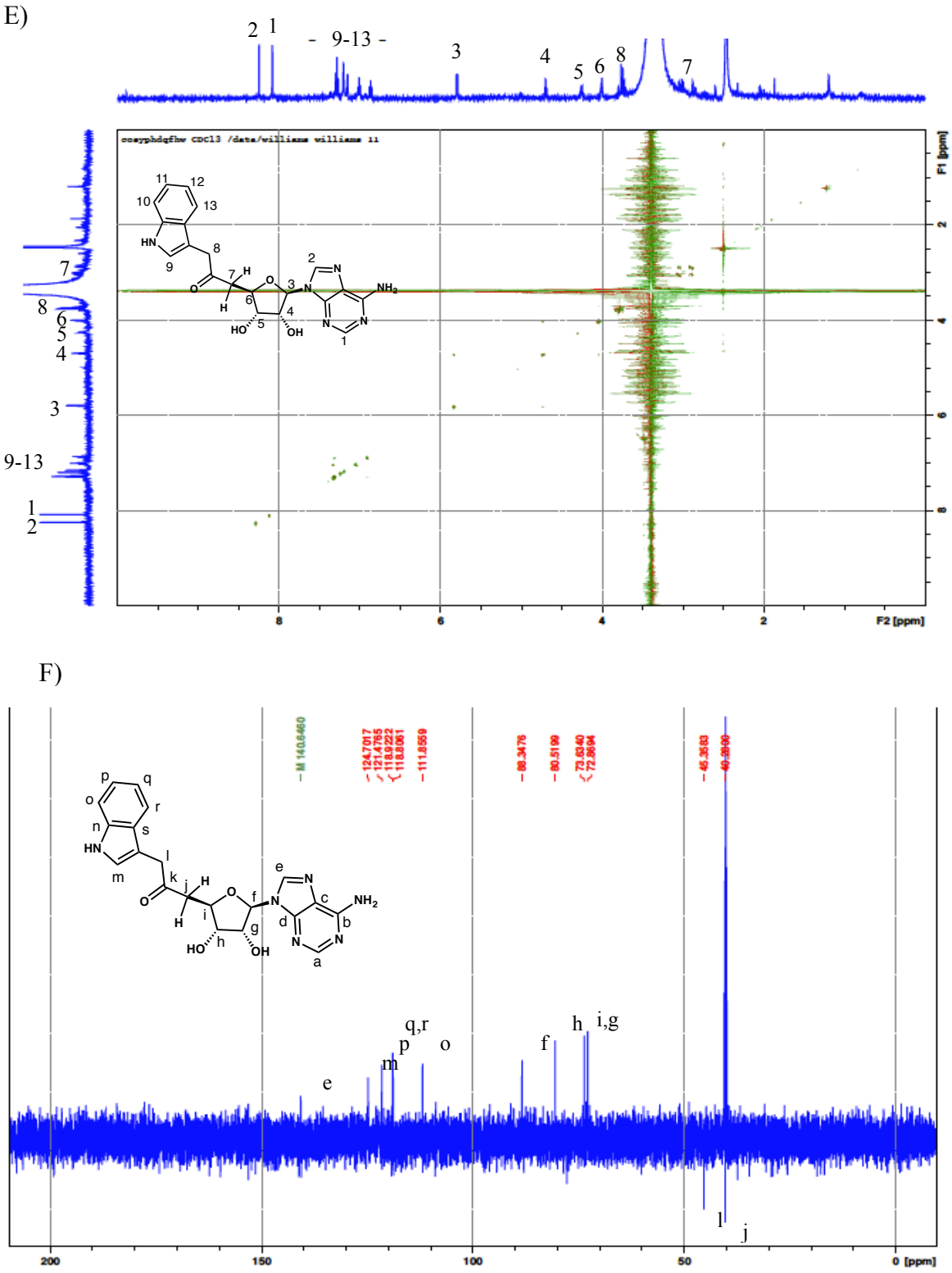


Figure 5.15 continued.

5.2.6 Reaction of NosL Y90A in absence of a substrate

The Y90F and Y90A NosL variants, in the absence of L-tryptophan (**13**) or its analogs, catalyzed the conversion of SAM to a new product, as indicated by HPLC analysis (Figure 5.16). UV-Vis analysis indicated the presence of the adenosine chromophore. LC-MS analysis gave an $[M-H]^-$ = 282 Da consistent with structure **173**. This was confirmed by derivatization with iodoacetamide (**174**), which resulted in a new species with $[M-H]^-$ = 339 Da consistent with **175**. Low levels of compound **173** were also observed with NosL WT as well as when FldA/FldR system was used.

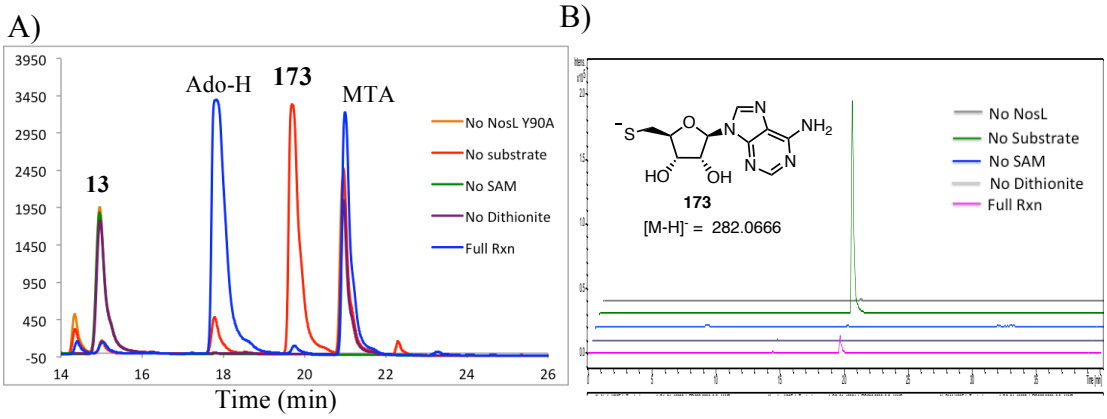


Figure 5.16: Analysis of the formation of 5'-deoxy-5'-thioadenosine (**173**). A) HPLC analysis at 260 nm of the NosL Y90A reaction with L-tryptophan (**13**). A new SAM-derived product is formed, in the absence of the amino acid substrate, that elutes at 20 min. B) LC-MS analysis showing EIC $[M-H]^-$ = 282.0 Da suggesting that this compound is **173**. C) MS of **173**. D) Derivatization of **173** using iodoacetamide (**174**) generates adduct **175.** E) LC-MS analysis showing EIC $[M-H]^-$ = 339.0 Da confirming formation of **175**. The signal of EIC $[M-H]^-$ = 282.0 Da disappeared after iodoacetamide treatment. E) MS of derivatized 5'-deoxy-5'-thioadenosine (**175**).

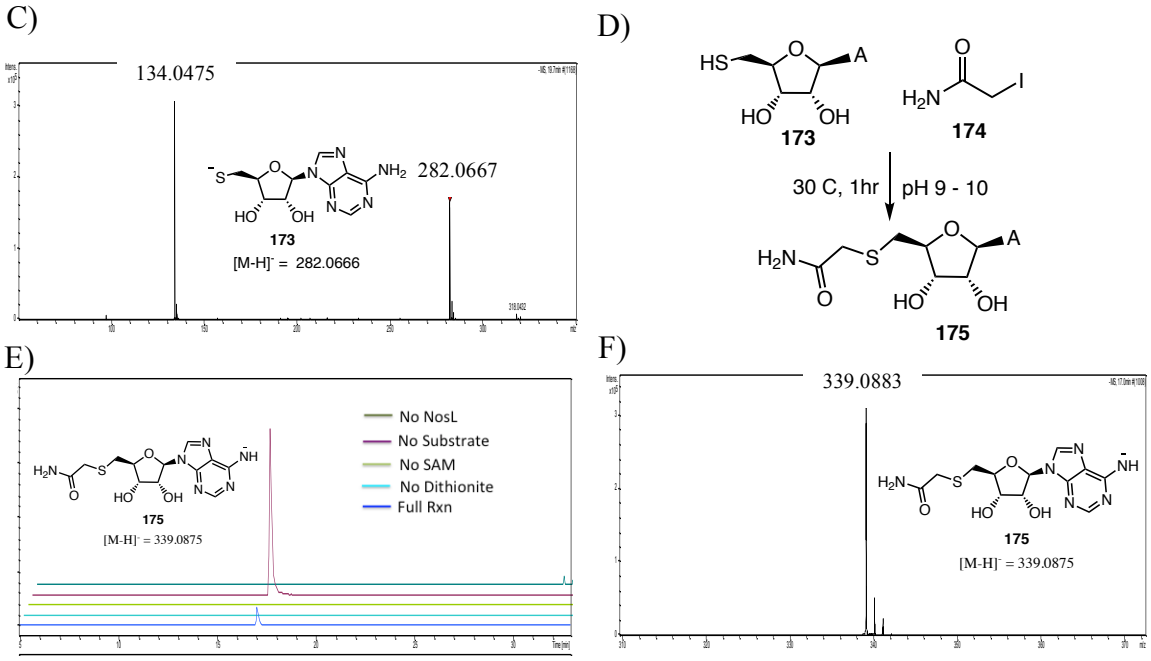


Figure 5.16 continued.

A mechanistic proposal for the formation of 5′-deoxy-5′-thioadenosine (**173**) is shown in Figure 5.17 (path a). In this proposal, the 5′-deoxyadenosyl radical (**6**) reacts with one of the sulfurs in the oxidized cluster to form **6a,** which decomposes to give **173**.[108] In NosL, formation of **173** is observed in higher amounts in Y90A and Y90F mutants when compared to WT. This suggests that residues like Y90 in the active site might act as H-atom donors to the 5'-deoxyadenosyl radical (**6**) during the uncoupled reaction. Recently, interaction of **6** with [4Fe-4S] was demonstrated using EPR.[109]

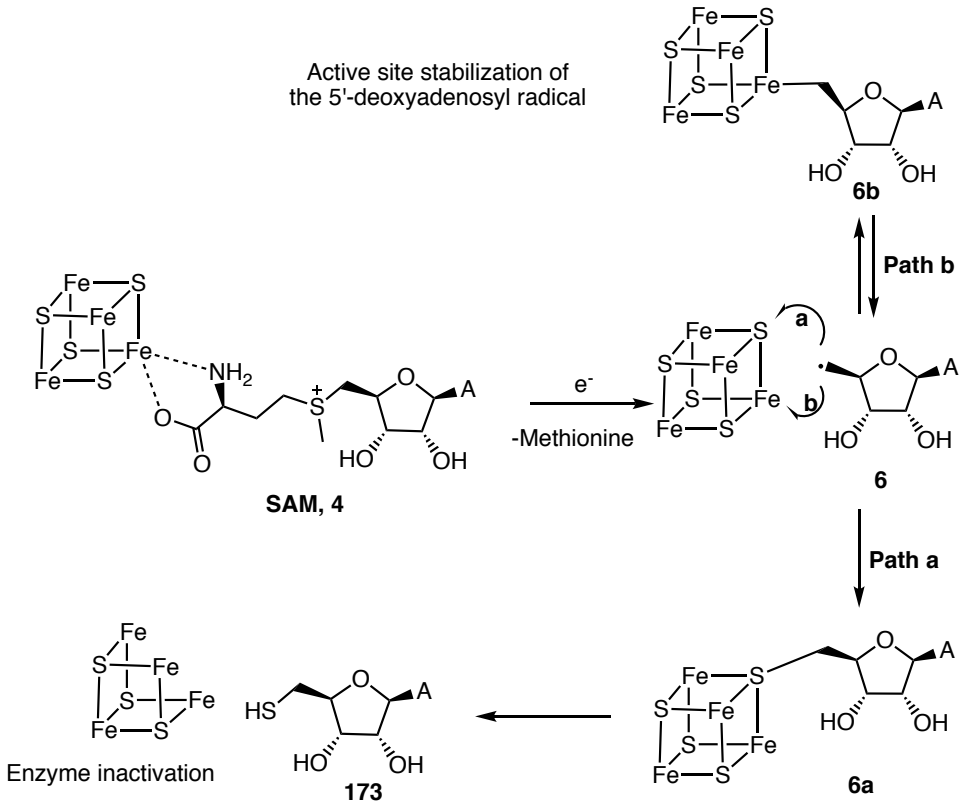
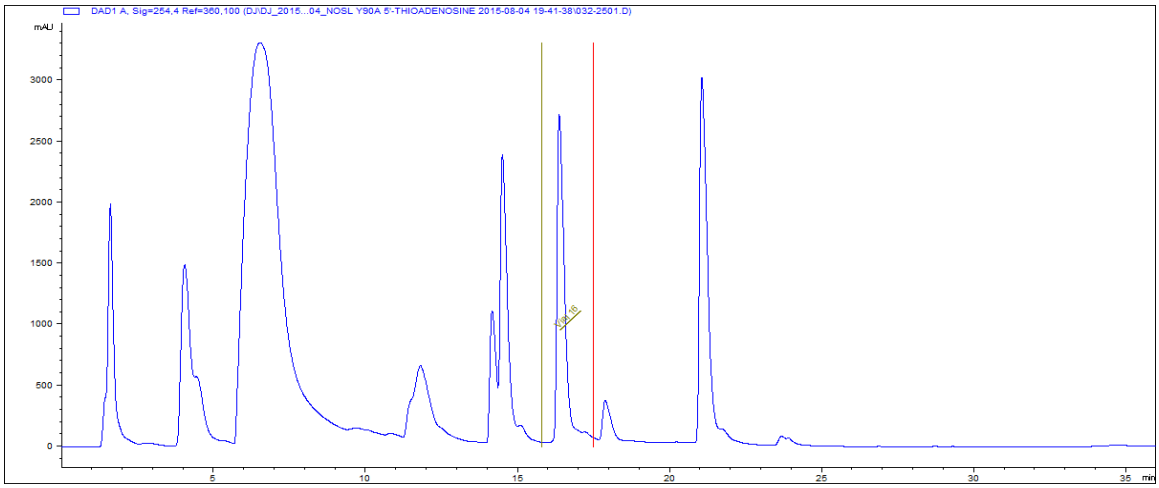


Figure 5.17: Mechanistic proposal for the formation of 5′-deoxy-5′-thioadenosine (**173**) involving C-S bond formation with the [4Fe-4S] cluster.

The iodoacetamide derivatized product **175** was purified by HPLC and its structure was confirmed by NMR analysis (Figure 5.18).

A)



Time (min)

B)

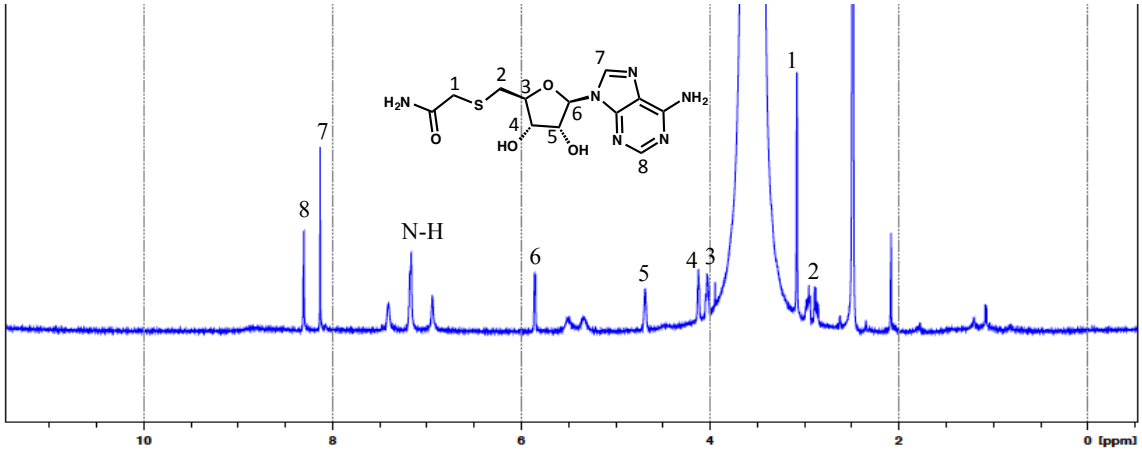


Figure 5.18: NMR characterization of the iodoacetamide derivative of 5'-deoxy-5'-thioadenosine **175**. A) HPLC analysis at 260 nm of the iodoacetamide derivatized NosL Y90A derived product. The signal that was collected is highlighted between the red and the green vertical lines. B) ^{1}H-NMR of the collected product **175**. C) ^{13}C-NMR of **175**. D) ^{1}H-^{1}H COSY-NMR of compound **175**.

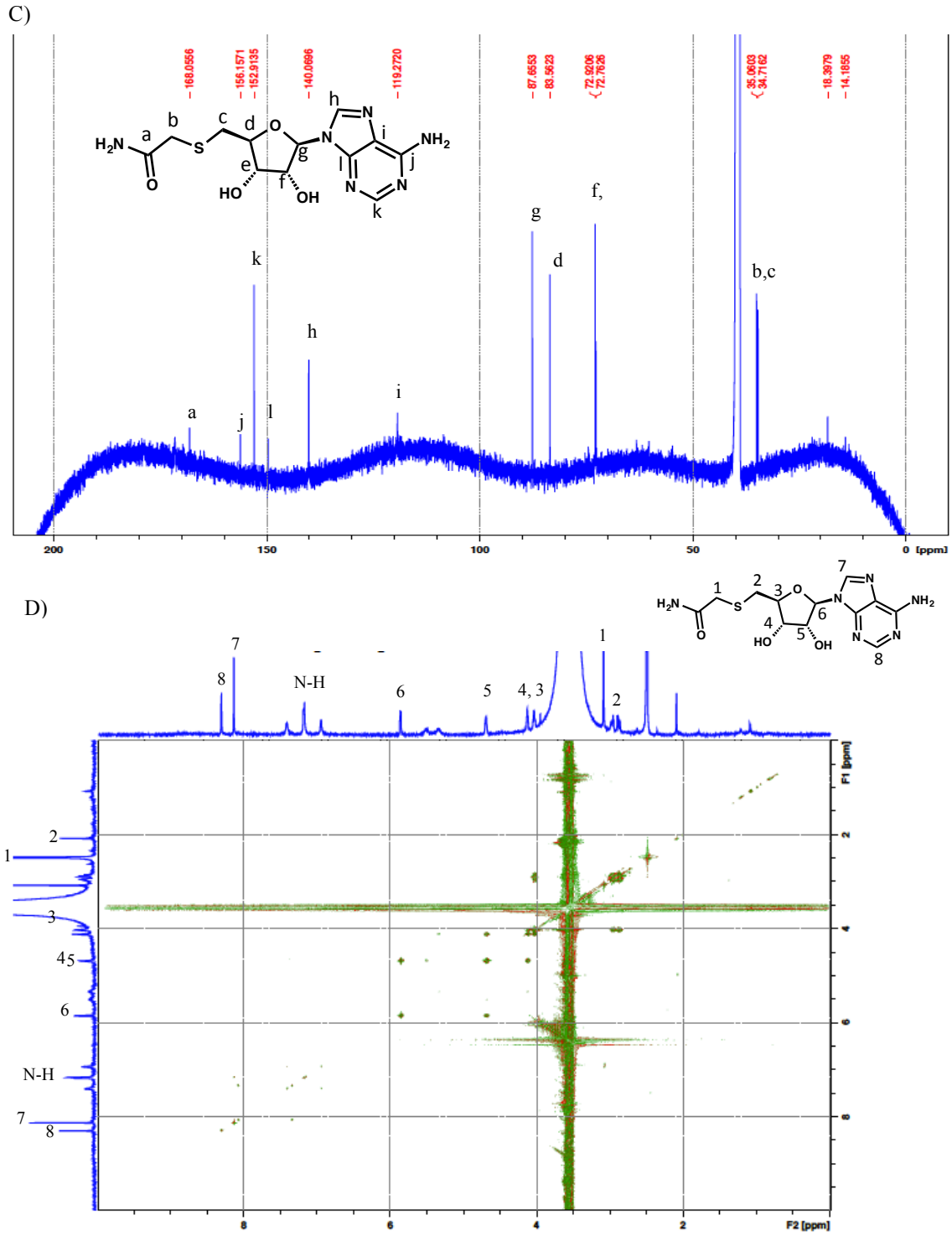


Figure 5.18 continued.

5.2.7 Reaction of NosL with L-tryptophanamide (176)

Enzyme-generated radical intermediates are sequestered from the buffer and generally cannot be trapped by exogenous reagents such as spin traps or hydrogen atom donors. Dithionite is emerging as an exception to this and sulfinic acid products have been reported for the spore photoproduct lyase,[110] Dph2[72] and DesII systems,[111] suggesting that **179** can diffuse into active sites and trap radical intermediates. This hypothesis was tested for the NosL catalyzed reaction.

LC-MS analysis of the NosL-catalyzed reaction of L-tryptophanamide (**176**) demonstrated the formation of a new product (Figure 5.19) having a mass consistent with sulfuramidite **180** resulting from the trapping of **176** by **179**.

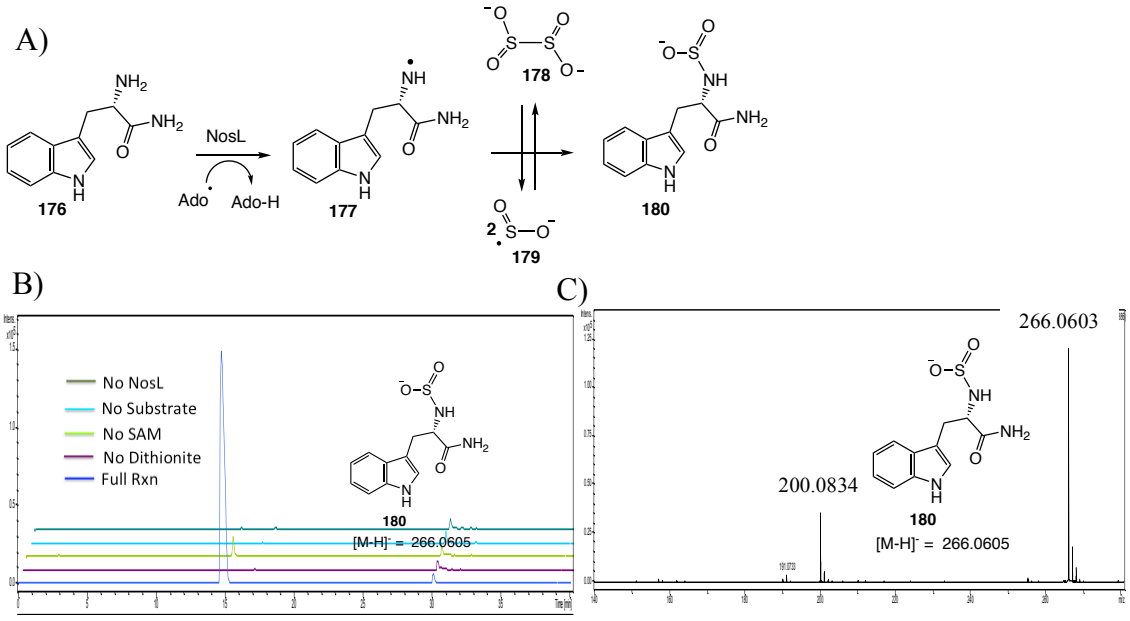


Figure 5.19: NosL-catalyzed formation of sulfinic acids. A) Generation of the sulfuramidite **180** with L-tryptophanamide (**176**). B) LC-MS analysis showing EIC $[M-H]^-$ = 266.0 Da suggesting that the new product is the sulfuramidite **180**. C) MS of the sulfuramidite **180**.

Similar adducts were observed with several other tryptophan analogs (N-methyl-L-tryptophanamide, N-ethyl-L-tryptophanamide, L-tryptophan methyl ester, L-tryptophan ethyl ester methyl-D,L-indole-3-lactate, ethyl-D,L-indole-3-lactate). In addition, 5′-deoxyadenosyl-5′-sulfinic acid (**187**) was observed in the absence of substrate and with enhanced levels in the presence of L-tryptophan ethyl ester. Summary of different dithionite adducts observed during investigation of the NosL catalyzed reaction are shown in Figure 5.20.

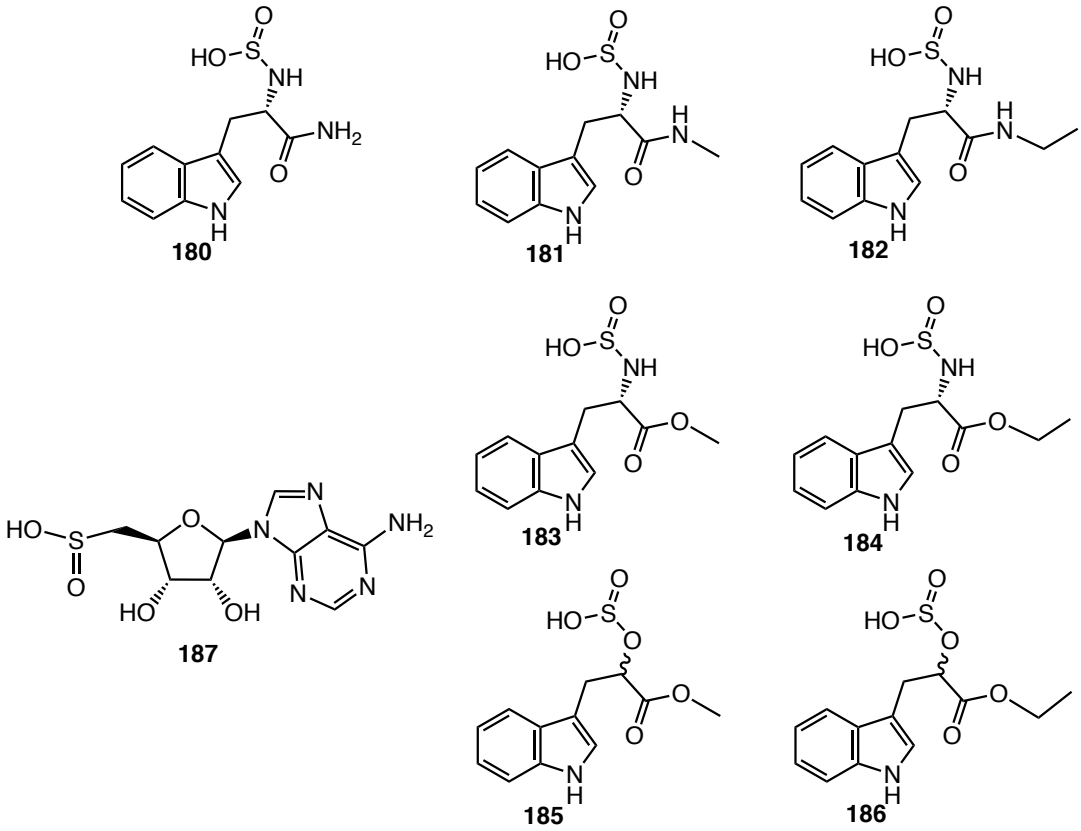


Figure 5.20: Summary of the dithionite adducts observed during the NosL-catalyzed reaction with the corresponding substrate analogs.

5.3 Conclusion

In this section we explored the reactivity of 5′-deoxyadenosyl radical (**6**) with NosL. We have demonstrated that N_{α}-alkylated tryptophan analogs are substrates for NosL and that the alkyl group (R) can alter the regiochemistry of the initial hydrogen atom abstraction from the amino group (R=H) to the alkyl group (R=Methyl) or to the Cα (R=cyclopropyl). Additionally, we have also demonstrated that the 5′-deoxyadenosyl radical (**6**) can abstract a H-atom from alcoholic O-H group if it is present in its vicinity.

We have also demonstrated that replacing the substrate amine with a ketone or an alkene changes the 5′-deoxyadenosyl radical chemistry from hydrogen atom abstraction to double bond addition and we propose that this strategy may be useful for the design of bisubstrate inhibitors of radical SAM enzymes. Such inhibitors would be of use for structural studies and as lead compounds for antibiotic development.

Additionally, we have also demonstrated that the 5′-deoxyadenosyl radical (**6**) can add to one of the sulfurs in the [4Fe-4S] cluster. C-S bond formation leading to **173** would result in enzyme inactivation and may be one of the contributing factors to the low turnover of enzymes in this superfamily.

Various sulfuramidites were trapped when the NosL reaction was performed in presence of dithionite. This suggests that dithionite has potential for the trapping of intermediates at the active sites of most radical SAM enzymes. Preliminary experiments with other radical SAM enzymes like MoaA,[112] BzaF,[113] MqnC,[114] and MqnE[106] suggest that many of these findings will have some generality in radical SAM enzymology.

5.4 Experimental procedures

Materials: All chemicals were purchased from Sigma-Aldrich unless specified otherwise. LB medium (Lennox formulation) was from EMD Millipore. Kanamycin was bought from Teknova and IPTG was obtained from Lab scientific Inc. Chloramphenicol was from Fisher Scientific. HPLC and LC-MS solvents were purchased from EMD and were used without further purification. Bradford reagent was purchased from ThermoFisher Scientific. Histrap column was obtained from GE healthcare. Econo-Pack 10DG desalting columns were purchased from Bio-Rad. D_2O was purchased from Cambridge Isotope Laboratories. Large cultures were grown and overexpressed in 2.5L baffled ultra yield flasks from Thomson Instrument Company.

Overexpression and purification of NosL, mutants of NosL, FldA and FldR: Gene of *nosl* from *Streptomyces actuosus* was used for all the studies. The overexpression and purification of all enzymes were performed as previously described in section 4.6.

Enzymatic reaction conditions for NosL: All enzymatic reactions of NosL and its mutants were carried out in an anaerobic chamber containing 95-97% nitrogen and 3-5% hydrogen. A typical enzymatic reaction was performed in 100 mM phosphate buffer, pH7.5 containing NosL (75-150 μM), dithionite (750 μM), L-tryptophan / substrate analog (400-500 μM) and SAM (600 μM). When assays were performed in the presence of the flavodoxin system, NosL was desalted using a Bio-Spin 6 column (Bio-Rad) into 100 mM phosphate buffer, pH 7.5. The components of the enzymatic reaction were

NosL (75-100 μM), FldA (60 μM), FldR (40 μM), NADPH (750 μM), L-tryptophan / substrate analogue (500 μM) and SAM (600 μM). The enzymatic reactions were incubated at room temperature for 2-3 h. Protein was removed by ultrafiltration using 10 kda cut-off filters (VWR) and the reaction mixture was further analyzed by HPLC and LC-MS.

HPLC and LC-MS parameters: These parameters were kept identical to the ones previously described in section 4.6.

NMR analysis: NMR spectra of all the synthetic samples were recorded on Bruker Avance III 400 MHz instrument. NMR characterization of samples collected by HPLC was performed on Bruker Avance III 500 MHz instrument with H-C-N cryoprobe in 3 mm Wilmad labglass (328-PP-7) high precision NMR tubes.

Synthesis of N_{α}-cyclopropyltryptophan (152): Substrate analog **152** was designed to probe the ring opening of cyclopropyl group after generation of a nitrogen-centered radical during the NosL-catalyzed reaction. This compound was synthesized according to the scheme shown in Figure 5.21A. Briefly, in a dry round bottom flask, 51 mg (0.25 mmol) of indole-3-pyruvic acid (**111**) was added to 8 ml dry methanol under an argon atmosphere. This was followed by addition of 70 μl (1 mmol) of cyclopropylamine (**153**) and 65 mg (1 mmol) of sodium cyanoborohydride. The reaction mixture was stirred under an inert atmosphere at room temperature for ~2 days. The precipitate of **152** was collected by vacuum filtration. Yield ~60 mg (98%). The

compound was characterized using NMR and LC-MS. For synthesis of Cα-deuterated isotope **152a**, $NaBD_3CN$ was used for reduction instead of $NaBH_3CN$ described in the above discussed procedure.

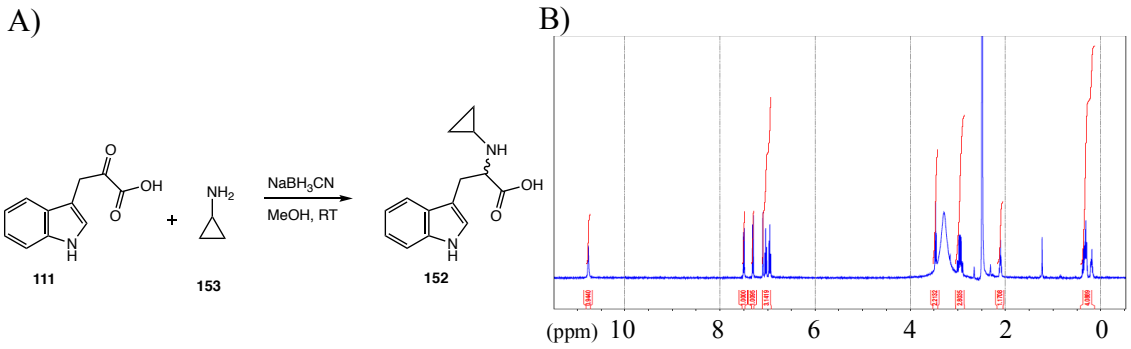


Figure 5.21: Synthesis of N_α-cyclopropyltryptophan (**152**). A) Synthetic scheme. B) ^{1}H-NMR of the synthesized **152**.

Synthesis of the methylene analog of tryptophan (165):[115] Compound **165** was synthesized according to the previously reported procedure as shown in Figure 5.22. Briefly, in a dry round bottom flask, add 2 ml dry tetrahydrofuran (THF) and 130 mg (6 mmol) Mg turnings. Under an inert atmosphere, add 650 mg (6 mmol) of ethyl bromide (**188**) dissolved in 2 ml THF. Stir on an ice bath for ~30 min after all Mg turnings have dissolved. Add dropwise solution of 590 mg (5 mmol) indole (**132**) dissolved in 2 ml THF and stir for 30 min. Dissolve 1 g (5.5 mmol) of methylbromomethacrylate (**189**) and add it drop wise. Remove the ice bath and stir at room temp for 30 min. Quench the reaction with water. Extract the aqueous phase with ethyl acetate and dry over $MgSO_4$. The pure ester **190** was obtained after chromatography (silica, hexane:ethyl acetate =

90:10). To hydrolyze the methyl ester of **190**, dissolve the purified product in 5 ml of 80% ethanol. Add 300 mg KOH and stir for 5 h at room temp. Then acidify the solution to pH 1 and extract the compound using diethyl ether. Purify **165** using flash chromatography (dichloromethane:methanol = 95:5). NMR was recorded in MeOD.

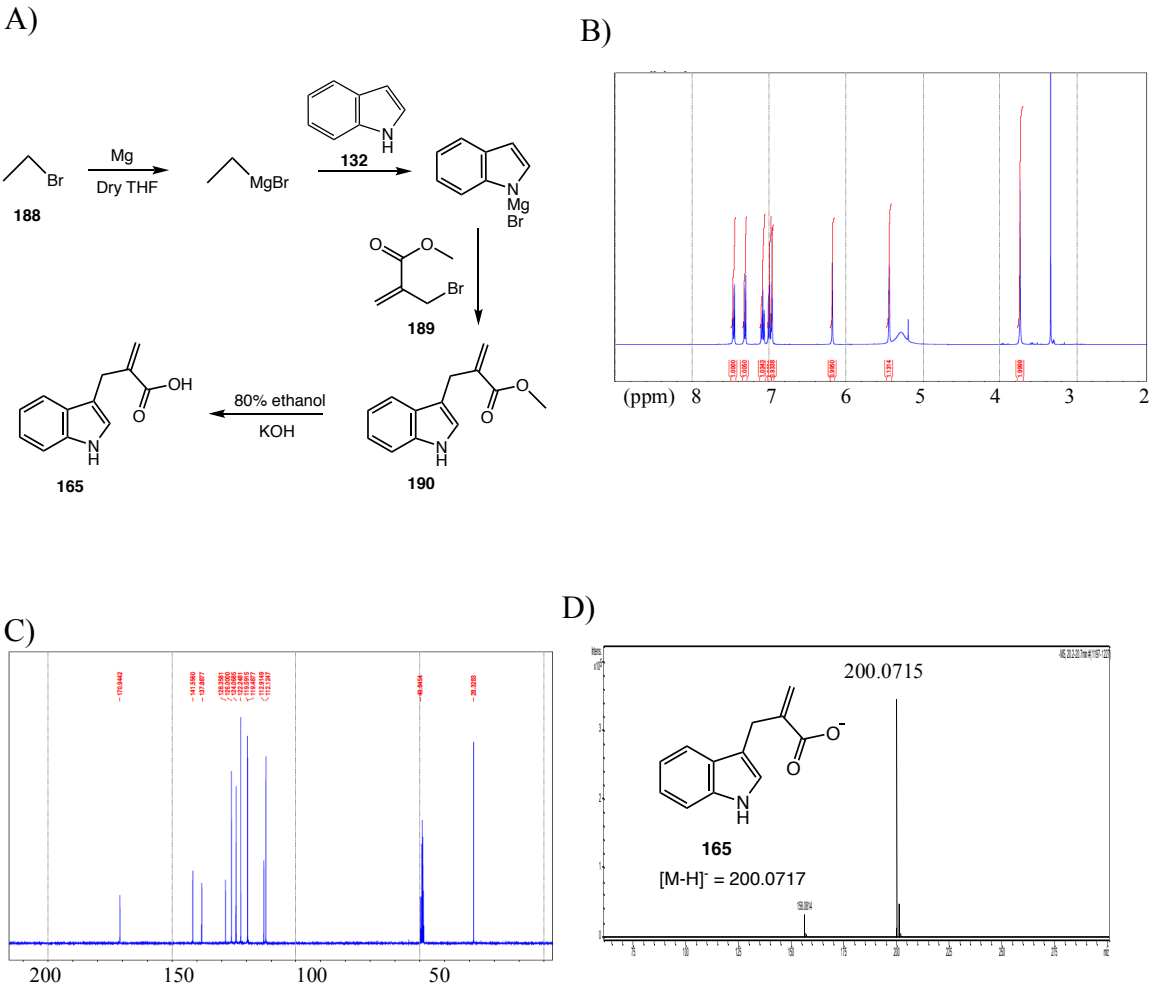


Figure 5.22: Synthesis and characterization of the methylene analog of tryptophan (**165**). A) Synthetic scheme. B) ^{1}H-NMR of **165**. C) ^{13}C-NMR of **165**. D) MS of **165** in negative mode.

6. SUMMARY AND OUTLOOK

The studies presented in this dissertation provide insights into the mechanistic work on three systems. Enzymes from a pathway that salvages L-cysteine from the S-substituted cysteines were reconstituted. Studies on the unique flavoenzyme CmoJ suggest that it catalyzes Pummerer-type rearrangement. Structural studies on Riboflavin Lyase (RcaE) illustrate that the orientations of the cofactor FMN and substrate riboflavin play a critical role for this novel motif in flavin enzymology, which utilizes superoxide radical to initiate chemistry on the substrate. Biochemical studies on Tryptophan Lyase (NosL) have highlighted the mobility of the 5′-deoxyadenosyl radical and advanced our understanding on the fragmentation-recombination reaction involved in nosiheptide biosynthetic pathway.

6.1 Cysteine salvage pathway

Amino acid cysteine and the metabolites derived from it play a vital role in all forms of life. Recently, a pathway was identified in *Bacillus subtilis* that salvages cysteine from S-substituted cysteines. In our study, we have elucidated the function of all five enzymes in this pathway. It involves a unique Pummerer-type rearrangement catalyzed by a flavoenzyme CmoJ. Though the pathway shows broad substrate specificity, our studies indicate that S-benzyl substituted cysteines might be the physiological substrates of the enzymes. Though there are few reports on biosynthesis of such molecules in plants, a comprehensive study will shed more light on their occurrence and distribution. Bioinformatic analysis shows that this pathway is present in

various microbes associated with plants. This study provides insights into a novel pathway in the largely unexplored world of rhizosphere microbiome.

6.2 Riboflavin catabolism

Strain *Microbacterium maritypicum* G10 was isolated from riboflavin rich environment and was shown to catabolize riboflavin to lumichrome and ribose. The key enzyme riboflavin lyase (RcaE) involved in this transformation expands the chemical repertoire of flavin superfamily. Structural studies on RcaE advance our mechanistic understanding on superoxide radical mediated transformation, a novel catalytic motif in flavin enzymology. Though we were unable to obtain the oxygen bound structure of the enzyme-cofactor-substrate complex, an oxygen molecule can be modeled near the C_{4a}-N_5 region of cofactor FMN and C_1' of the ribose chain in riboflavin (Figure 3.8B). This demonstrates that RcaE exhibits a steric control to catalyze this innovative superoxide radical mediated transformation. Future experiments will demonstrate the distribution of this motif.

The pathway discussed here utilizes the ribose from riboflavin as a carbon source and does not degrade the alloxazine ring of lumichrome. This suggests that other organisms may have evolved pathways to utilize lumichrome as a potential source of carbon as well as nitrogen. Additionally, this study sets the stage for further studies on alternate riboflavin catabolic pathways, which were previously reported but the strains were lost. Screening of bacterial strains collected from the riboflavin production plant (DSM nutritional products, Germany) is currently underway in Begley lab to explore these hypotheses on riboflavin catabolism.

6.3 Tryptophan Lyase (NosL)

Radical SAM enzymes catalyze some of the most complex rearrangement reactions discovered in the biological world. Tryptophan Lyase (NosL) from the aromatic amino acid lyases sub-family involves a fragmentation-recombination reaction to generate 3-methyl-2-indolic acid from L-tryptophan. We have identified cyanide as a byproduct of this reaction and demonstrated that the 5′-deoxyadenosyl radical abstracts H-atom from the amino group of L-tryptophan. Mutagenesis studies show that Arg323 is important for controlling the regiochemistry of the β-scission reaction. Further studies with L-tryptophan thiocarboxylate and 2'-halo-L-tryptophan support the Cα-C(O) cleavage after H-atom abstraction. A revised mechanistic proposal consistent with all observations discussed and recently observed EPR spectroscopy data has been illustrated in section 4.4.11 (Figure 4.34). This study has uncovered an unanticipated connection between NosL and HydG – a radical SAM enzyme that forms cyanide and carbon monoxide from tyrosine during the biosynthesis of the metallo cluster of the [Fe-Fe]-hydrogenase.

While investigating the mechanism of the NosL catalyzed reaction, we explored the reactivity of the 5′-deoxyadenosyl radical. We have demonstrated that the regiochemistry of H-atom abstraction can be altered, the chemistry can be changed from H-atom abstraction to double bond addition, attack of 5′-deoxyadenosyl radical to [4Fe-4S] leads to enzyme inactivation and radical intermediates can be trapped by dithionite. This study highlights a plethora of transformations possible by the incredible 5′-deoxyadenosyl radical, which is a common intermediate in most radical SAM enzymes.

REFERENCES

1. Richter, M., *Natural Product Reports* **2013,** *30* (10), 1324-1345.

2. Begley, T. P.; Kinsland, C.; Taylor, S.; Tandon, M.; Nicewonger, R.; Wu, M.; Chiu, H.-J.; Kelleher, N.; Campobasso, N.; Zhang, Y., In *Biosynthesis: Polyketides and Vitamins*, Leeper, F. J.; Vederas, J. C., Eds. Springer Berlin, Heidelberg: **1998**, 93-142.

3. Northrop-Clewes, C. A.; Thurnham, D. I., *Annals of Nutrition and Metabolism* **2012,** *61* (3), 224-230.

4. Bacher, A.; Eberhardt, S.; Eisenreich, W.; Fischer, M.; Herz, S.; Illarionov, B.; Kis, K.; Richter, G., *Vitamins & Hormones* **2001,** *61*, 1-49.

5. Massey, V., *Biochemical Society Transactions* **2000**, 28 (4), 283-296.

6. Powers, H. J., *The American Journal of Clinical Nutrition* **2003,** *77* (6), 1352-1360.

7. Fontecave, M.; Atta, M.; Mulliez, E., *Trends in Biochemical Sciences* **2004,** *29* (5), 243-249.

8. Cantoni, G., *Journal of Biological Chemistry* **1951,** *189*, 203-216.

9. Lu, S. C., *The International Journal of Biochemistry & Cell Biology* **2000,** *32* (4), 391-395.

10. Sofia, H. J.; Chen, G.; Hetzler, B. G.; Reyes-Spindola, J. F.; Miller, N. E., *Nucleic Acids Research* **2001,** *29* (5), 1097-1106.

11. Gerlt, J. A.; Bouvier, J. T.; Davidson, D. B.; Imker, H. J.; Sadkhin, B.; Slater, D. R.; Whalen, K. L., *Biochimica et Biophysica Acta (BBA)-Proteins and Proteomics* **2015,** *1854* (8), 1019-1037.

12. Frey, P. A.; Hegeman, A. D.; Ruzicka, F. J., *Critical Reviews in Biochemistry and Molecular Biology* **2008,** *43* (1), 63-88.

13. Mehta, A. P.; Abdelwahed, S. H.; Mahanta, N.; Fedoseyenko, D.; Philmus, B.; Cooper, L. E.; Liu, Y.; Jhulki, I.; Ealick, S. E.; Begley, T. P., *Journal of Biological Chemistry* **2015,** *290* (7), 3980-3986.

14. Chan, C. M.; Danchin, A.; Marlière, P.; Sekowska, A., *Environmental Microbiology* **2014,** *16* (1), 101-117.

15. Foster, J. W., *Journal of Bacteriology* **1944,** *47* (1), 27-41.

16. Foster, J. W., *Journal of Bacteriology* **1944,** *48* (1), 97-111.

17. Yang, C. S.; McCormick, D. B., *Biochimica et Biophysica Acta (BBA)-Enzymology* **1967,** *132* (2), 511-513.

18. Mander, L.; Liu, H.-W., *Comprehensive Natural Products II: Chemistry and Biology*. Elsevier Science: **2010**, Vol. 1.

19. Xu, H.; Chakrabarty, Y.; Philmus, B.; Mehta, A. P.; Bhandari, D.; Hohmann, H.-P.; Begley, T. P., *Journal of Biological Chemistry* **2016,** *291* (45), 23506-23515.

20. Yu, Y.; Duan, L.; Zhang, Q.; Liao, R.; Ding, Y.; Pan, H.; Wendt-Pienkowski, E.; Tang, G.; Shen, B.; Liu, W., *ACS Chemical Biology* **2009,** *4* (10), 855-864.

21. Zhang, Q.; Li, Y.; Chen, D.; Yu, Y.; Duan, L.; Shen, B.; Liu, W., *Nature chemical biology* **2011,** *7* (3), 154-160.

22. Nicolet, Y.; Zeppieri, L.; Amara, P.; Fontecilla-Camps, J. C., *Angewandte Chemie International Edition* **2014,** *53* (44), 11840-11844.

23. Frigaard, N.-U.; Dahl, C., *Advances in Microbial Physiology* **2008,** *54*, 103-200.

24. Stipanuk, M. H., *Annual Review of Nutrition* **1986,** *6* (1), 179-209.

25. Py, B.; Barras, F., *Nature Reviews. Microbiology* **2010,** *8* (6), 436-446.

26. Nuesch, J.; Heim, J.; Treichler, H., *Annual Reviews in Microbiology* **1987,** *41* (1), 51-75.

27. Cullis, C.; Hirschler, M., *Atmospheric Environment (1967)* **1980,** *14* (11), 1263-1278.

28. Bollinger, J. M.; Price, J. C.; Hoffart, L. M.; Barr, E. W.; Krebs, C., *European Journal of Inorganic Chemistry* **2005,** *2005* (21), 4245-4254.

29. Kahnert, A.; Vermeij, P.; Wietek, C.; James, P.; Leisinger, T.; Kertesz, M. A., *Journal of Bacteriology* **2000,** *182* (10), 2869-2878.

30. Gallagher, J. R.; Olson, E. S.; Stanley, D. C., *FEMS Microbiology Letters* **1993,** *107* (1), 31-35.

31. Riddles, P. W.; Blakeley, R. L.; Zerner, B., *Methods in Enzymology* **1983,** *91*, 49-60.

32. Deuerling, E.; Schulze-Specking, A.; Tomoyasu, T.; Mogk, A.; Bukau, B., *Nature* **1999,** *400* (6745), 693-696.

33. Simpson, R. T.; Vallee, B. L., *Inorganic Chemistry* **1969,** *8* (5), 1185-1186.

34. Chinn, P. C.; Pigiet, V.; Fahey, R. C., *Analytical Biochemistry* **1986,** *159* (1), 143-149.

35. Spyrou, G.; Haggård-Ljungquist, E.; Krook, M.; Jörnvall, H.; Nilsson, E.; Reichard, P., *Journal of Bacteriology* **1991,** *173* (12), 3673-3679.

36. Weissbach, H.; Resnick, L.; Brot, N., *Biochimica et Biophysica Acta (BBA)-Proteins and Proteomics* **2005,** *1703* (2), 203-212.

37. Kubec, R.; Musah, R. A., *Phytochemistry* **2001,** *58* (6), 981-985.

38. Wang, Z., In *Comprehensive Organic Name Reactions and Reagents*, John Wiley & Sons, Inc.: Hoboken, **2009**.

39. Conte, M. L.; Carroll, K. S., In *Oxidative Stress and Redox Regulation*, Springer: **2013**, 1-42.

40. Klomsiri, C.; Nelson, K. J.; Bechtold, E.; Soito, L.; Johnson, L. C.; Lowther, W. T.; Ryu, S.-E.; King, S. B.; Furdui, C. M.; Poole, L. B., *Methods in Enzymology* **2010,** *473*, 77-94.

41. Gupta, V.; Carroll, K. S., *Biochimica et Biophysica Acta (BBA)-General Subjects* **2014,** *1840* (2), 847-875.

42. Byrn, M.; Calvin, M., *Journal of the American Chemical Society* **1966,** *88* (9), 1916-1922.

43. Schummer, C.; Delhomme, O.; Appenzeller, B. M.; Wennig, R.; Millet, M., *Talanta* **2009,** *77* (4), 1473-1482.

44. Holmgren, A., *Journal of Biological Chemistry* **1989,** *264* (24), 13963-13966.

45. Ellis, H. R.; Poole, L. B., *Biochemistry* **1997,** *36* (48), 15013-15018.

46. Miron, T.; Rabinkov, A.; Mirelman, D.; Weiner, L.; Wilchek, M., *Analytical Biochemistry* **1998,** *265* (2), 317-325.

47. Song, H.; Hu, W.; Naowarojna, N.; Her, A. S.; Wang, S.; Desai, R.; Qin, L.; Chen, X.; Liu, P., *Scientific Reports* **2015,** *5,* 11870.

48. Adak, S.; Begley, T. P., *Journal of the American Chemical Society* **2016,** *138* (20), 6424-6426.

49. Robbins, J. M.; Ellis, H. R., *Biochemistry* **2012,** *51* (32), 6378-6387.

50. Tschantz, W. R.; Digits, J. A.; Pyun, H.-J.; Coates, R. M.; Casey, P. J., *Journal of Biological Chemistry* **2001,** *276* (4), 2321-2324.

51. Zerbs, S.; Korajczyk, P. J.; Noirot, P. H.; Collart, F. R., *Protein Science* **2017,** *26* (4), 784-795.

52. Pedrolli, D. B.; Kühm, C.; Sévin, D. C.; Vockenhuber, M. P.; Sauer, U.; Suess, B.; Mack, M., *Proceedings of the National Academy of Sciences* **2015,** *112* (45), 14054-14059.

53. Nishihara, K.; Kanemori, M.; Yanagi, H.; Yura, T., *Applied and Environmental Microbiology* **2000,** *66* (3), 884-889.

54. Marasco, E. K.; Schmidt-Dannert, C., *ChemBioChem* **2008,** *9* (9), 1450-1461.

55. Bray, H.; James, S. P.; Thorpe, W., *Biochemical Journal* **1958,** *70* (4), 570-579.

56. Baxter, R. L.; McGregor, C. J.; Thomson, G. A.; Scott, A. I., *Journal of the Chemical Society, Perkin Transactions 1* **1985,** (0), 369-372.

57. Meese, C. O.; Specht, D.; Hofmann, U., *Archiv der Pharmazie* **1990,** *323* (12), 957-965.

58. Mukherjee, T.; McCulloch, K.; Ealick, S.; Begley, T., *Comprehensive Natural Products II, Chemistry and Biology* **2010,** *7*, 649-674.

59. Begley, T. P.; Chatterjee, A.; Hanes, J. W.; Hazra, A.; Ealick, S. E., *Current Opinion in Chemical Biology* **2008,** *12* (2), 118-125.

60. Kersters, K.; Ludwig, W.; Vancanneyt, M.; De Vos, P.; Gillis, M.; Schleifer, K.-H., *Systematic and Applied Microbiology* **1996,** *19* (4), 465-477.

61. Marchena, M.; Gil, M.; Martín, C.; Organero, J. A.; Sanchez, F.; Douhal, A., *The Journal of Physical Chemistry B* **2011,** *115* (10), 2424-2435.

62. Massey, V., *Journal of Biological Chemistry* **1994,** *269* (36), 22459-22462.

63. Hassan, Y. I.; Lepp, D.; He, J.; Zhou, T., *Genome Announcements* **2014,** *2* (5), e00994-14.

64. Broderick, J. B.; Duffus, B. R.; Duschene, K. S.; Shepard, E. M., *Chemical Reviews* **2014,** *114* (8), 4229-4317.

65. Pryor, W., *Free Radicals in Biology*. Elsevier: **2012**; Vol. 6.

66. Meunier, B.; de Visser, S. P.; Shaik, S., *Chemical Reviews* **2004,** *104* (9), 3947-3980.

67. Tinberg, C. E.; Lippard, S. J., *Accounts of Chemical Research* **2011,** *44* (4), 280-288.

68. Reichard, P.; Ehrenberg, A., *Science* **1983,** *221* (4610), 514-519.

69. Fraaije, M. W.; Mattevi, A., *Trends in Biochemical Sciences* **2000,** *25* (3), 126-132.

70. Banerjee, R., *Chemical Reviews* **2003,** *103* (6), 2083-2094.

71. Hernandez, D.; François, P.; Farinelli, L.; Østerås, M.; Schrenzel, J., *Genome Research* **2008,** *18* (5), 802-809.

72. Zhang, Y.; Zhu, X.; Torelli, A. T.; Lee, M.; Dzikovski, B.; Koralewski, R. M.; Wang, E.; Freed, J.; Krebs, C.; Ealick, S. E., *Nature* **2010,** *465* (7300), 891-896.

73. Demick, J. M.; Lanzilotta, W. N., *Biochemistry* **2011,** *50* (4), 440-442.

74. Mancia, F.; Keep, N. H.; Nakagawa, A.; Leadlay, P. F.; McSweeney, S.; Rasmussen, B.; Diat, O.; Evans, P. R., *Structure* **1996,** *4* (3), 339-350.

75. Benazet, F.; Cartier, M.; Florent, J.; Godard, C.; Jung, G.; Lunel, J.; Mancy, D.; Pascal, C.; Renaut, J.; Tarridec, P., *Experientia* **1980,** *36* (4), 414-416.

76. Haste, N. M.; Thienphrapa, W.; Tran, D. N.; Loesgen, S.; Sun, P.; Nam, S.-J.; Jensen, P. R.; Fenical, W.; Sakoulas, G.; Nizet, V., *The Journal of Antibiotics* **2012,** *65* (12), 593-598.

77. Ortega, M. A.; van der Donk, W. A., *Cell Chemical Biology* **2016,** *23* (1), 31-44.

78. Kriek, M.; Martins, F.; Challand, M. R.; Croft, A.; Roach, P. L., *Angewandte Chemie International Edition* **2007,** *46* (48), 9223-9226.

79. Duffus, B. R.; Ghose, S.; Peters, J. W.; Broderick, J. B., *Journal of the American Chemical Society* **2014,** *136* (38), 13086-13089.

80. Bhandari, D. M.; Xu, H.; Nicolet, Y.; Fontecilla-Camps, J. C.; Begley, T. P., *Biochemistry* **2015,** *54* (31), 4767-4769.

81. Ji, X.; Li, Y.; Ding, W.; Zhang, Q., *Angewandte Chemie International Edition* **2015,** *54* (31), 9021-9024.

82. Philmus, B.; Decamps, L.; Berteau, O.; Begley, T. P., *Journal of the American Chemical Society* **2015,** *137* (16), 5406-5413.

83. Ji, X.; Liu, W.-Q.; Yuan, S.; Yin, Y.; Ding, W.; Zhang, Q., *Chemical Communications* **2016,** *52* (69), 10555-10558.

84. Wagner, P. J.; Lindstrom, M. J.; Sedon, J. H.; Ward, D. R., *Journal of the American Chemical Society* **1981,** *103* (13), 3842-3849.

85. Joshi, S.; Mahanta, N.; Fedoseyenko, D.; Williams, H.; Begley, T. P., *Journal of the American Chemical Society* **2017,** *139* (32), 10952-10955.

86. Chatterjee, A.; Hazra, A. B.; Abdelwahed, S.; Hilmey, D. G.; Begley, T. P., *Angewandte Chemie International Edition* **2010,** *49* (46), 8653-8656.

87. Tracqui, A.; Raul, J.; Geraut, A.; Berthelon, L.; Ludes, B., *Journal of Analytical Toxicology* **2002,** *26* (3), 144-148.

88. Kawasaki, H.; Bauerle, R.; Zon, G.; Ahmed, S.; Miles, E., *Journal of Biological Chemistry* **1987,** *262* (22), 10678-10683.

89. Cancilla, D. A.; Hee, S. S. Q., *Journal of Chromatography A* **1992,** *627* (1-2), 1-16.

90. Sicoli, G.; Mouesca, J.-M.; Zeppieri, L.; Amara, P.; Martin, L.; Barra, A. L.; Fontecilla-Camps, J. C.; Gambarelli, S.; Nicolet, Y., *Science* **2016,** *351* (6279), 1320-1323.

91. Driesener, R. C.; Challand, M. R.; McGlynn, S. E.; Shepard, E. M.; Boyd, E. S.; Broderick, J. B.; Peters, J. W.; Roach, P. L., *Angewandte Chemie International Edition* **2010,** *49* (9), 1687-1690.

92. Pagnier, A.; Martin, L.; Zeppieri, L.; Nicolet, Y.; Fontecilla-Camps, J. C., *Proceedings of the National Academy of Sciences* **2016,** *113* (1), 104-109.

93. Kuchenreuther, J. M.; Myers, W. K.; Suess, D. L.; Stich, T. A.; Pelmenschikov, V.; Shiigi, S. A.; Cramer, S. P.; Swartz, J. R.; Britt, R. D.; George, S. J., *Science* **2014,** *343* (6169), 424-427.

94. O'Leary, M. H., *Annual Review of Plant Physiology* **1982,** *33* (1), 297-315.

95. Wellner, D.; Meister, A., *Journal of Biological Chemistry* **1960,** *235* (7), 2013-2018.

96. Walsh, C. T.; Schonbrunn, A.; Abeles, R. H., *Journal of Biological Chemistry* **1971,** *246* (22), 6855-6866.

97. Mali, S. M.; Jadhav, S. V.; Gopi, H. N., *Chemical Communications* **2012,** *48* (56), 7085-7087.

98. Phillips, R. S.; Cohen, L. A., *Tetrahedron Letters* **1983,** *24* (50), 5555-5558.

99. Banerjee, R.; Ragsdale, S. W., *Annual Review of Biochemistry* **2003,** *72* (1), 209-247.

100. Nam, W., *Accounts of Chemical Research* **2007,** *40* (7), 522-531.

101. Hioe, J.; Zipse, H., *Chemistry-A European Journal* **2012,** *18* (51), 16463-16472.

102. Horitani, M.; Byer, A. S.; Shisler, K. A.; Chandra, T.; Broderick, J. B.; Hoffman, B. M., *Journal of the American Chemical Society* **2015,** *137* (22), 7111-7121.

103. Imai, K.; Watanabe, Y., *Analytica Chimica Acta* **1981,** *130* (2), 377-383.

104. Bhandari, D. M.; Fedoseyenko, D.; Begley, T. P., *Journal of the American Chemical Society* **2016,** *138* (50), 16184-16187.

105. Ji, X.; Li, Y.; Xie, L.; Lu, H.; Ding, W.; Zhang, Q., *Angewandte Chemie International Edition* **2016,** *55* (39), 11845-11848.

106. Mahanta, N.; Fedoseyenko, D.; Dairi, T.; Begley, T. P., *Journal of the American Chemical Society* **2013,** *135* (41), 15318-15321.

107. Wagner, A. V.; Demand, J.; Schilling, G.; Pils, T.; Knappe, J., *Biochemical and Biophysical Research Communications* **1999,** *254* (2), 306-310.

108. Farrar, C. E.; Jarrett, J. T., *Biochemistry* **2009,** *48* (11), 2448-2458.

109. Horitani, M.; Shisler, K.; Broderick, W. E.; Hutcheson, R. U.; Duschene, K. S.; Marts, A. R.; Hoffman, B. M.; Broderick, J. B., *Science* **2016,** *352* (6287), 822-825.

110. Chandor-Proust, A.; Berteau, O.; Douki, T.; Gasparutto, D.; Ollagnier-de-Choudens, S.; Fontecave, M.; Atta, M., *Journal of Biological Chemistry* **2008,** *283* (52), 36361-36368.

111. Ko, Y.; Ruszczycky, M. W.; Choi, S. H.; Liu, H. w., *Angewandte Chemie* **2015,** *127* (3), 874-877.

112. Mehta, A. P.; Hanes, J. W.; Abdelwahed, S. H.; Hilmey, D. G.; Hänzelmann, P.; Begley, T. P., *Biochemistry* **2013,** *52* (7), 1134-1136.

113. Mehta, A. P.; Abdelwahed, S. H.; Fenwick, M. K.; Hazra, A. B.; Taga, M. E.; Zhang, Y.; Ealick, S. E.; Begley, T. P., *Journal of the American Chemical Society* **2015,** *137* (33), 10444-10447.

114. Cooper, L. E.; Fedoseyenko, D.; Abdelwahed, S. H.; Kim, S.-H.; Dairi, T.; Begley, T. P., *Biochemistry* **2013,** *52* (27), 4592-4594.

115. Buddoo, S.; Rousseau, A. L.; Gordon, G. E., Cargill Inc., *U.S. Patent No. 8367847*, **2013**.